LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

UCRL-JRNL-207451

On The Reproducibility of Seasonal Land-surface Climate

T. J. Phillips

October 25, 2004

Journal of Hydrometeorology

Disclaimer

On the Reproducibility of Seasonal Land-surface Climate

Thomas J. Phillips

Program for Climate Model Diagnosis and Intercomparison
Lawrence Livermore National Laboratory, L-103, Livermore, California 94551, USA
email: phillips14@llnl.gov; phone: +1-925-422-0072; fax: +1-925-422-7675

Submission in revised form to the *Journal of Hydrometeorology*
April, 2005

ABSTRACT

The sensitivity of the continental seasonal climate to initial conditions is estimated from an ensemble of decadal simulations of an atmospheric general circulation model with the same specifations of radiative forcings and monthly ocean boundary conditions, but with different initial states of atmosphere and land. As measures of the "reproducibility" of continental climate for different initial conditions, spatio-temporal correlations are computed across paired realizations of eleven model land-surface variables in which the seasonal cycle is either included or excluded--the former case being pertinent to climate simulation, and the latter to seasonal anomaly prediction.

It is found that the land-surface variables which include the seasonal cycle are impacted only marginally by changes in initial conditions; moreover, their seasonal climatologies exhibit high spatial reproducibility. In contrast, the reproducibility of a seasonal land-surface anomaly is generally low, although it is substantially higher in the Tropics; its spatial reproducibility also markedly fluctuates in tandem with warm and cold phases of the El Niño/Southern Oscillation. However, the overall degree of reproducibility depends strongly on the particular land-surface anomaly considered. It is also shown that the predictability of a land-surface anomaly implied by its reproducibility statistics is consistent with what is inferred from more conventional predictability metrics. Implications of these results for climate model intercomparison projects and for operational forecasts of seasonal continental climate also are elaborated.

1. Introduction

While the implications of initial-condition sensitivity (i.e. "chaos") for the prediction of climate have been recognized for decades (e.g. Lorenz 1964, Leith 1973), practical investigation of this problem with general circulation models (GCMs) did not begin until the mid-1980s (Chervin 1986, Zwiers 1987). Only in the past several years, moreover, has computer performance advanced sufficiently to permit routine generation of experimental ensembles for studies of climate prediction on a range of time scales.

Cubasch et al. (1994), for example, showed that the multi-decadal climate simulated by a synchronously coupled ocean-atmosphere model is sensitive to the initial conditions, an issue also revisited by Delworth and Knutson (2000) for transient global-warming experiments. Barnett (1995) noted that chaos is also a factor in "two-tier" experiments wherein the ocean and atmosphere models are asynchronously coupled, and Hansen et al. (1997) documented its effects at annual to decadal time scales in coupled climate-change simulations involving a mixed-layer ocean.

Many other investigators (a partial list includes Palmer and Anderson 1994, Dix and Hunt 1995, Kumar and Hoerling 1995, Stern and Miyakoda 1995, Bengtsson et al. 1996, Zwiers 1996, Carson 1998, Zwiers and Kharin 1998, Wang and Zwiers 1999, Koster et al. 2000, Shukla et al. 2000, Dirmeyer 2001, and Dirmeyer et al. 2001) have demonstrated that, even when ocean boundary conditions are prescribed, simulations of climate also are chaotic on seasonal time scales. However, many of these seasonal predictability studies have focussed on the dynamics of the free atmosphere rather than on the processes at the land surface, the chief locus of human interaction with the climate system. Even when continental processes were considered (e.g. Dix and Hunt 1995, Stern and Miyakoda 1995, Wang and Zwiers 1999, Koster et al. 2000, Dirmeyer 2001), only a limited number

of land-surface variables (e.g. precipitation and soil moisture) were investigated. In light of the potentially enormous societal ramifications, a more comprehensive assessment of the predictability of seasonal land-surface climate is clearly necessary.

It also should be noted that the initial-condition sensitivity of continental climate is germane for climate modeling activities that lie outside the sphere of predictability studies *per se*. The most clear-cut examples include modeling experiments where chaos may confound attempts to attribute model variations in simulated seasonal climates to differences in coupled land-surface schemes (e.g. Henderson-Sellers et al. 1996; Polcher 2000; Dirmeyer et al. 2001; Koster et al. 2002, 2004); but such issues are also pertinent for virtually all model intercomparison projects that consider land-surface processes.

Thus, there is an overarching need to identify when and where continental climate simulations are likely to be most reproducible in the context of varying model initial conditions. From the perspective of model intercomparison studies, it is the "reproducibility" of the entire climatic state--not just the anomalous departures from the climatic mean--that is of concern. In this conceptual framework, therefore, "reproducibility" subsumes "predictability", which usually pertains only to the potential for predicting climate *anomalies*.

Climate reproducibility exhibits both temporal and spatial aspects. For a surface variable V(x,y,t) to possess high spatial reproducibility at time t, its (x,y) pattern must remain similar across realizations R_1, R_2, ..., R_n having somewhat different initial conditions . For the surface variable to possess high temporal reproducibility, the trajectories V(t) in time of each (x,y) point of variable V(x,y,t) must also remain similar across different realizations (see Figure 1).

Different climate variables may display more or less temporal and spatial reproducibility. For example, because precipitation is a fine-scale bipolar process that can be switched "on" or "off" depending on small variations in atmospheric humidity, vertical motion, or thermodynamic stability, its climatic mean (e.g. seasonal average) might be expected to exhibit relatively low temporal reproducibility, especially in regions where the precipitation "on/off " action is most pronounced. In contrast, climatic-mean precipitation might be expected to show relatively high spatial reproducibility, since certain regions (e.g. tropical rain forests or subtropical deserts) are consistently wet or dry; but this spatial reproducibility also might fluctuate seasonally or interannually, for example with changing SST patterns. Temporal reproducibility thus may vary spatially, and spatial reproducibility temporally. In order to really analyze such phenomena, however, it is necessary to go beyond intuitive conceptions of reproducibility, and to define appropriate statistical metrics (see Section 4).

In this paper, the spatio-temporal reproducibility of a diverse collection of seasonal land-surface variables are estimated from an ensemble of decadal climate simulations implemented by a version of the European Centre for Medium-Range Weather Forecasts (ECMWF) GCM with prescribed ocean boundary conditions. The reproducibility of each land-surface variable when its climatological seasonal cycle is included or excluded is separately considered, since these configurations are pertinent to climate model intercomparison or to seasonal prediction, respectively.

Section 2 summarizes the salient features of the ECMWF model version and of the experimental design, while Section 3 compares the simulated interannual variability with observational analyses, in order to evaluate the model's suitability for this study. Section 4 defines spatio-temporal reproducibility metrics, computes these statistics for 11 seasonal land-surface vari-

ables and corresponding seasonal anomalies, and presents results for selective cases. Section 5 demonstrates that the predictability implied by the reproducibility statistics of each seasonal land-surface anomaly is consistent with more conventional predictability metrics. Finally, Section 6 elaborates the implications of this study for both climate model intercomparison and seasonal prediction.

2. Experimental Design

a. Model features

This study employed cycle 36 of the ECMWF atmospheric GCM, a model version that was in operational use during the early 1990s (see features summary by Phillips 1994). In this model, the atmospheric primitive equations are represented as truncated series of spherical harmonic basis functions, where nonlinear terms and many physical parameterizations are calculated on a Gaussian grid and then transformed to the spectral space. In this study's model experiments, the horizontal resolution was triangularly truncated at wave number 42 (corresponding to a 64 x 128 Gaussian grid), and vertical differences were expressed on 19 unevenly spaced levels in hybrid sigma-pressure coordinates. For a surface pressure of 1000 hPa, the lowest prognostic level was at 996 hPa, and 5 levels were below 800 hPa.

The prognostic equations were solved semi-implicitly at a 30-minute time step, but with radiative fluxes (after Morcrette 1991) and clouds (after Slingo 1987) updated only every 3 hours. Convection was formulated as in Tiedtke (1989). Dissipation was implemented as biharmonic horizontal diffusion and stability-dependent second-order vertical diffusion, and orographic gravity-wave drag was parameterized after Miller et al. (1989). In the surface layer, fluxes of momentum, heat, and moisture were treated as in Louis (1979), with Miller et al. (1992) modifications of stability-dependent drag coefficients to enhance ocean surface evaporation for calm conditions.

Land-surface processes were parameterized after Blondin (1989). A vegetation canopy occupying a variable fraction of each grid box intercepted a portion of the incident precipitation, which subsequently evaporated at the potential rate. Transpiration of soil moisture was regulated by a canopy resistance that depended on local radiation and moisture stresses, but not on vegetation type. Evaporation and sublimation from the bare-soil and snow-covered fractions of each grid box also were treated.

Land temperature and moisture profiles were predicted by diffusion equations in top and middle soil layers (thickness 0.07 m and 0.42 m, respectively), but were prescribed from monthly climatologies in an underlying deep layer. When the moisture capacity of the top soil layer (.02 m) was exceeded, surface runoff occurred; base flow resulted from overflow of the middle soil layer (moisture capacity 0.12 m). Some inherent shortcomings of this land-surface scheme for simulating continental climate also have been noted (e.g. Betts et al. 1993). These problems motivated revisions of the land-surface parameterizations in subsequent versions of the ECMWF model (Viterbo and Beljaars 1995, Beljaars et al. 1996).

Several potential sources of chaotic behavior may be identified *a priori* in such a land-surface scheme. First, the initial specification of soil temperature/moisture and snow cover can impact the scheme's equilibrium state (Yang et al. 1995). Because the fraction of precipitation that is intercepted by the vegetation canopy evaporates quickly, the variability of surface evaporation is directed toward higher frequencies than would be the case if a canopy were not represented (Scott et al. 1997). Moreover, by coupling this scheme to an atmospheric model that is also sensitive to initial conditions, the overall chaotic behavior of the system is accentuated.

Thus, a general challenge to analyzing the effects of a coupled land-surface scheme is that

the strength and spatial pattern of the land-atmosphere coupling is likely to be model-specific (e.g. see results from the GLACE ensemble experiments reported by Koster et al. 2002, 2004). This is also the case when different land-surface schemes are coupled to the same atmospheric model, in that changes may occur in the interannual variability of the respective ensemble-mean climates that appear to be related to variations in land-atmosphere coupling (e.g. Maynard and Polcher 2003). The results of the present study therefore are likely to depend to some degree on the particular biases and land-atmosphere coupling characteristics of the ECMWF cycle 36 model.

b. Ensemble experiments

The available computer resources allowed an ensemble of 6 decadal climate simulations to be generated, where each realization had common ocean boundary conditions obtained from Atmospheric Model Intercomparison Project (AMIP) specifications of observed monthly sea surface temperatures (SSTs) and sea ice extents for the period 1 January 1979 to 31 December 1988 (Gates 1992, Gates et al. 1999). (These boundary conditions were updated daily by linear interpolation of contiguous monthly fields.) In addition, AMIP values of the solar constant (1365 $W m^{-2}$) and of the carbon dioxide concentration (345 ppm) were specified.

The nominal start time of each realization was also the same (0000 UMT on 1 January 1979), but the initial conditions of atmospheric and continental variables differed. For the first realization of the period 1979 to 1988, these initial conditions were specified from the ECMWF observational analysis for 0000 UMT on 15 January 1979 (taken as the effective date of the monthly mean January 1979 AMIP SST and sea ice extents). For each of the 5 companion realizations of the same time period, a different initial state that also was representative of January climate was imposed. For example, the initial conditions (at 0000 UMT on 1 January 1979) for the second realization

were specified to be the same as the state at the last time step (i.e., at 23:30 UMT on 31 December 1988) of the first realization, and so on. Given the omnipresence of chaotic nonlinear interactions, the specification of somewhat different initial atmospheric states for the 6 simulations was sufficient to produce statistically independent realizations of decadal climate. (Essentially the same result would be achieved by generating a single 60-year run with SSTs recycled every 10 years, and then appropriately treating the model output as a sequence of 6 independent decadal realizations.)

3. The Simulated Surface Climate

a. Model output data

For each decadal realization, time series of monthly averages of eleven AMIP "standard output variables" (Gates 1992) were computed, providing a reasonably comprehensive portrait of the simulated surface climate (Table 1). Time series also were calculated of the seasonal averages of these variables, i.e. for each sample of March-April-May (MAM), June-July-August (JJA), September-October-November (SON), and December-January-February (DJF) climates. (Owing to the simulation start date on 1 January 1979, only 9 samples of DJF climates were available, resulting in a time series of 39 seasonal samples per decadal realization.)

As expected, these seasonal averages were less sensitive to initial conditions than were the monthly means, since longer time-averaging filters out more of the climate noise engendered by the chaotic daily weather systems (Leith 1973, Barnston 1994). Averaging over periods longer than a season would further decrease the climate noise (Ebisuzaki 1995, Chen and van den Dool 1997), but at the expense of reducing the number of samples available from the decadal simulations. Hence, detailed analysis of the simulation ensemble was limited to seasonal time scales.

b. Validation of model variability

Accurate estimation of the reproducibility of seasonal land-surface climate from such decadal-scale ensemble experiments demands that the model adequately simulate the observed interannual climate variability (e.g. Kumar et al. 1996, Sperber and Palmer 1996, Liang et al. 1997). Because there can be nontrivial differences in simulated interannual variability among ensemble members (e.g. local peak values range between ~ 50 to 150 $(\text{mm day}^{-1})^2$ for precipitation) the best available estimate of model interannual variability is provided by the ensemble-mean value of this statistic across the six climate realizations. Observations reflect only one realization of climate variability, but by using several data sets based on different observing systems and methodologies as validation references, current uncertainties in the observed climate variability can be inferred. If the model is performing adequately, the ensemble-mean interannual variability of a simulated variable should largely fall within these estimated observational uncertainties.

As a basic reality-check of the ECMWF cycle 36 model, the ensemble-mean interannual variabilities of three key seasonal surface variables--JJA and DJF mean sea-level pressure (MSLP), surface air temperature, and precipitation--are compared against estimates from different observationally based data sets. For reasons that will become more apparent in Section 5b, the NCEP-NCAR reanalysis (Kalnay et al. 1996) is chosen as one such validation reference because it provides a comprehensive and self-consistent picture of 1979-1988 surface climate variability that is independent of the ECMWF model. Appropriate alternative reference data sets also are chosen. For example, in validating the interannual variability of model seasonal MSLP, the NCEP-R2 and the ERA15 reanalyses are also selected as validation references (Figure 2). The NCEP-R2 (Kanamitsu et al. 2002) reanalysis

improves upon certain assimilation techniques of NCEP-NCAR, while the ERA15 reanalysis (Gibson et al. 1997) is based on an ECMWF model of more recent vintage than cycle 36 which also includes improved parameterizations of land-surface processes (Viterbo and Beljaars 1995).

The reference estimates of MSLP interannual variability all show close agreement in both JJA and DJF, with some differences occurring at high latitudes (Figure 2). The model-simulated interannual variability of MSLP (Figures 2a and 2e) also generally agrees well with these validation references. The chief discrepancies in JJA (Figure 2a) include somewhat too little model variability near Antarctica and too much near Greenland. In DJF (Figure 2e), the model variability near Antarctica is again too low, and that associated with the Icelandic Low is centered somewhat too far east.

For surface air temperature (Figure 3), the chosen validation references include the NCEP-NCAR and NCEP-R2 reanalyses and the Climatic Research Unit (CRU) land-only (excluding Greenland and Antarctica) data set based on station observations (New et al. 2002). The model closely agrees with both reanalyses' indications of a center of variability in the Tropical East Pacific associated with El Niño-Southern Oscillation (ENSO) phenomena. Over land, the reference estimates are in fair agreement, with centers of surface temperature variability over North America and Eurasia in JJA and DJF, but differing somewhat in magnitude and placement. The model simulation of temperature variability over land appears to fall within the implied observational uncertainties in JJA (Figure 3a), but seems somewhat too low over the Northern continents in DJF (Figure 3e).

Because the NCEP-NCAR precipitation product is influenced substantially by the parameterizations of the associated analysis model (Kalnay et al. 1996), blended satellite and gauge estimates of precipitation variability provided by the Global Precipitation Climatology Project (GPCP,

Huffman and Bolvin 2004) and the solely gauge-based terrestrial estimates of the CRU data set (New et al. 2002) are used as supplemental validation references (Figure 4). The reference estimates display considerable similarity at large scales, but with regional-scale differences in magnitude and pattern, notably in the Tropics in both seasons and over the Northern continents in JJA.

The model simulation of the large-scale patterns of precipitation variability matches these reference estimates fairly well. However, the magnitude of the simulated variability seems somewhat too large, especially over the tropical oceans and the Northern continents in JJA, implying that model convection may be too vigorous. The discrepancies between the simulated precipitation variability and the estimates provided by the NCEP-NCAR reanalysis probably also reflect somewhat the different convective schemes of the ECMWF cycle 36 and NCEP reanalysis models (Tiedtke 1989, Pan and Wu 1994).

In summary, the ECMWF cycle 36 model produces a generally credible large-scale simulation (relative to the observational uncertainties implied by the chosen validation reference data) of the interannual variability of key surface variables. There are, however, some apparent shortcomings in simulating details of this variability in certain regions and seasons. Perhaps the most troubling model deficiency is the anomalous rendering of precipitation over Northern continents in JJA, since the simulated reproducibility of land-surface variables is likely to be most susceptible to model biases during boreal summer, when the coupling of soil moisture and precipitation is comparatively strong (e.g. Koster et al. 2002, 2004).

4. Reproducibility Metrics for Seasonal Surface Variables

In this section, spatio-temporal reproducibility metrics are defined and then are applied to the 6 realizations of surface variables obtained from the ECMWF model simulations.

a. Spatio-temporal measures of reproducibility

To make intuitive conceptions of spatio-temporal reproducibility more precise, appropriate statistical measures need to be defined (e.g. Santer 1988, Dix and Hunt 1995). For example, to quantify the temporal reproducibility of a model surface variable, zero-lag temporal cross-correlations $r^{l,m}(i,j)$ may be computed between each independent pair (l,m) of realizations that are available from the ensemble of size n = 6 members at every grid point (i,j) . That is, for surface variable V

$$r^{l,m}(i,j) = \sum_{k}^{K} \{V^{l}(i,j,k) - \overline{V}^{l}(i,j)\} * \{V^{m}(i,j,k) - \overline{V}^{m}(i,j)\} / \{K * \sigma_t^{l}(i,j) * \sigma_t^{m}(i,j)\} \qquad \begin{matrix} l = 1, 2, ..., n = 6 \\ m = 1, 2, ..., n = 6 \end{matrix}$$

where $\overline{V}$ denotes the temporal mean of V, σ_t is its temporal standard deviation from $\overline{V}$, and the summation is over all K = 39 seasonal samples k in a decadal realization.

As a measure of the spatial reproducibility for each season k of surface variable V, contemporaneous spatial (pattern) correlations $s(k)$ also may be computed between each pair (l,m) of realizations:

$$s^{l,m}(k) = \sum_{i,j}^{I \times J} \{V^{l}(i,j,k) - [V^{l}(k)]\} * \{V^{m}(i,j,k) - [V^{m}(k)]\} / \{(IxJ) * \sigma_S^{l}(k) * \sigma_S^{m}(k)\}$$

Here, the summation is over the total number IxJ of surface grid points (i,j) for the chosen 64x128 Gaussian grid (or ~ 2500 grid points when only land surfaces are considered), [V] denotes the spatial average of V, and σ_s is its corresponding spatial standard deviation about [V].

For an ensemble of n statistically independent members, the number N of independent real-

ization pairs (l,m) is

$$N = n!/\{2!(n-2)!\}$$

In this study, for example, the n = 6 model realizations yield N = 15 samples of cross-correlations *r* and *s*. Thus, as an overall measure of reproducibility, the ensemble mean (denoted < >) and intraensemble scatter (denoted Δ) of the cross-correlations *r* and *s* are calculated over these N samples. In the case of temporal cross-correlation fields *r*(i,j), for example, these ensemble statistics are computed over N realization pairs (l,m) as follows:

$$\langle r(i,j) \rangle = \sum_{l,m}^{N} r^{l,m}(i,j)/ N$$

and

$$\Delta r(i,j) = \left\{ \sum_{l,m}^{N} [r^{l,m}(i,j) - \langle r(i,j) \rangle]^2 / N \right\}^{1/2}$$

Analogous calculations pertain for determining the ensemble mean $\langle s(k) \rangle$ and intraensemble scatter $\Delta s(k)$ for each season k from N samples of spatial cross-correlations *s*(k,l,m).

b. Reproducibility of the model surface variables

Model surface variables that include the seasonal cycle (relevant for climate model intercomparison projects) are analyzed first. In this context, *r* and *s* essentially measure the reproducibility of each variable's seasonal cycle, which accounts for most of its spatio-temporal variability about the decadal climatological mean.

1) Temporal reproducibility of seasonal variables

Figure 5 shows maps of $\langle r(i,j) \rangle$ and $\Delta r(i,j)$ for the simulated seasonal precipitation, a variable

that exhibits comparatively low temporal reproducibility. These maps display pronounced latitudinal asymmetries, in that the highest mean cross-correlations $<r(i,j)>$ occur in the equatorial Tropics. Here ocean-land contrasts also are evident, with correlations greater than 0.9 covering broad swaths of the equatorial Atlantic and Pacific oceans, but only portions of the adjacent continental areas in Amazonia, equatorial Africa, and southeast Asia. At most extratropical locations where chaotic dynamics prevail, $<r(i,j)>$ is less than 0.7, implying that a single realization l of extratropical seasonal precipitation typically explains less than half the temporal variance of another realization m. In general, the intraensemble scatter $\Delta r(i,j)$ also is substantially higher outside the deep Tropics, notably in subtropical regions dominated by high pressure, where precipitation is sparse or episodic.

An area-weighted average R of the cross-correlation field $r(i,j)$ of model variable V over all land points provides an aggregate measure of its continental temporal reproducibility. In Figure 6a, the ensemble mean $<R>$ and intraensemble scatter ΔR of area-average R across the N=15 samples (l,m) of $r(i,j)$ are shown for each seasonal-mean model variable. The values of <R> are seen to range widely over different variables, with the ground and surface air temperatures (**tg** and **tas**) having the highest values (~1), and precipitation (**pr**) the lowest (~ 0.6). For each land-surface variable, the intraensemble scatter ΔR (gray bar) is much smaller than $<R>$ (dark bar), but ΔR does tend to increase as $<R>$ decreases. The statistics $<R>$ and ΔR therefore consistently depict the variations in aggregate temporal reproducibility displayed by model land-surface variables that include the seasonal cycle.

In Figure 6b, the $<R>$ and ΔR are displayed instead for *monthly* samples of the same land-surface variables, where the annual cycle is included. As noted previously, the monthly land-surface

variables all exhibit somewhat less temporal reproducibility than the corresponding seasonal variables (Figure 6a), owing to the comparatively greater sensitivity of the monthly statistics to varying initial conditions. This behavior is especially pronounced for land-surface precipitation (**pr**).

2) Spatial reproducibility of seasonal surface variables

The spatial reproducibility of simulated land-surface variables that include the seasonal cycle, as exemplified by the variation of the mean spatial cross-correlation $<s(k)>$ for precipitation (Figure 7), shows substantial variation with season k. Over land, $<s(k)>$ displays a pronounced sawtooth fluctuation each year (Figures 7b, 7c), with relative maxima repeatedly occurring in JJA and DJF and minima in MAM and SON, presumably because the continental precipitation patterns are less stable in the latter transition seasons. (The modest scatter $\Delta s(k)$ also shown in Figures 7b and 7c implies that such seasonal variation in $<s(k)>$ is consistently exhibited by the correlation samples.) Sawtooth fluctuations also characterize the $<s(k)>$ time series for global precipitation (Figure 7a), except in 1983 and 1987 when this pattern is disrupted, probably related to occurrences of El Niño in these years. It is also seen that the annual-average amplitude of the precipitation pattern correlations over tropical land points (Figure 7c) is larger than that for the all-land and global cases.

Similar seasonal variations in $<s(k)>$ are displayed by all other model land-surface variables except for soil moisture, which peaks in MAM and SON instead of JJA and DJF each year. Because soil moisture is a source of climate memory, it is physically reasonable for it to be in seasonal quadrature with precipitation and other model variables. This relationship also implies soil moisture's potential as a predictor of seasonal climate (e.g. Huang et al. 1996; Koster et al. 2000, 2002, 2004).

3) Spatial reproducibility of seasonal surface climatologies

Because model intercomparison projects often are limited to analysis of single realizations of simulated climatologies, it is relevant to consider the reproducibility of the ECMWF surface variables' climatological seasons τ (i.e., the decadal average MAM, JJA, SON, and DJF). This is accomplished by computing the mean and scatter among the spatial cross-correlations of the n=6 realizations of each seasonal climatology τ that were available for a surface variable V. For all variables and seasons, the resulting ensemble-mean correlations were ~ 1, with intraensemble scatter less than ~ 0.03. Hence, the intercomparison of single realizations of seasonal surface climatologies from a collection of atmospheric models *with common prescribed ocean boundary conditions and radiative forcings* seems to be a potentially meaningful exercise, at least on the continental-to-global scales where today's GCMs are thought to provide meaningful predictions.

c. Reproducibility of seasonal surface anomalies

From the vantage point of seasonal prediction, it is the reproducibility of a seasonal anomaly A (i.e. the departure of surface variable V from its climatological MAM, JJA, SON, or DJF mean) which is of interest. That is, V is predictable to the extent that A is reproducible--otherwise, the seasonal climatology offers the better prediction of V.

In the AMIP runs, however, the ECMWF model's ocean state is not prognostically determined, so that the chosen reproducibility measures allow only the inference of a seasonal variable's *potential predictability* (hereafter, PP), a presumed upper bound on forecast skill (Madden 1981). Thus, in order to infer the PP of seasonal surface variable V, the ten-year climatological seasonal cycle is removed from its time series, and then the cross-correlations *r* and *s* of the resulting seasonal anomaly A for each pair of realizations are computed. As summary measures, the ensem-

ble mean $<r(i,j)>$ and intraensemble scatter $\Delta r(i,j)$ statistics from the N = 15 samples of *r* and *s* also again are calculated.

1) Temporal reproducibility of seasonal anomalies

As expected, values of $<r(i,j)>$ for the model's seasonal surface anomalies are substantially less than for the corresponding surface variables with the seasonal cycle included. For example, maps of $<r(i,j)>$ and $\Delta r(i,j)$ for seasonal anomalies of MSLP (**psl**), which exhibit relatively high temporal reproducibility, are shown in Figure 8. Values of $<r(i,j)>$ in excess of 0.3 extend into the extratropics as far as western North America and northern Australia (Figure 8a), farther poleward than for any of the other model surface anomalies (not shown). The intraensemble scatter $\Delta r(i,j)$ for MSLP (Figure 8b) is also generally less than for other anomalies.

The overall temporal reproducibility of each land-surface anomaly is indicated by the ensemble mean $<R>$ and intraensemble scatter ΔR of area-weighted land-averages *R* of *r*(i,j), which are shown for all land points in Figure 9a, and for only tropical land points in Figure 9b. In both cases, anomalies of MSLP (**psl**), surface air temperature (**tas**), and ground temperature (**tg**) have the highest $<R>$; it is disconcerting, however, that anomalies of surface evaporation (**evs**), precipitation (**pr**), and soil moisture (**mrso**) display comparatively low <R>, in view of the socio-economic importance of these hydrological variables. For each surface anomaly, the tropical-land average $<R>$ is substantially higher, and its scatter ΔR lower, than the corresponding all-land statistics, which are degraded by the inclusion of highly chaotic extratropical points. Especially noteworthy are the tropical-land MSLP (**psl**) and surface air temperature (**tas**) anomalies, whose $<R>$ values exceed 0.40, with low scatter ΔR.

The fact that the model surface anomalies are generally less temporally reproducible in the

extratropics is probably a consequence of the baroclinic instabilities and associated turbulent eddies that are endemic there. In the Tropics, larger-scale barotropic dynamical modes predominate, which may account for the relatively high temporal reproducibility of the pressure and temperature anomalies. The low temporal reproducibility of tropical precipitation and turbulent surface fluxes may then be mainly a consequence of their inherently finer scales.

2) Spatial reproducibility of seasonal anomalies

In contrast, model precipitation anomalies generally show the greatest spatial reproducibility, as evinced by the time series of the ensemble-mean pattern correlation $<s(k)>$ (Figure 10). For instance, there are many seasons k when $<s(k)>$ for the global (land + ocean) precipitation anomalies is > 0.5 (Figure 10a). There also is relatively low scatter $\Delta s(k)$ in the time series of the global precipitation anomalies; however, the spatial reproducibility of the precipitation anomalies is substantially reduced over land (Figures 10a vs. 10b), especially in middle and high latitudes (Figures 10b vs. 10c). Other model surface anomalies also exhibit reduced spatial reproducibility over the continents, but to a lesser degree than precipitation. For example, the average amplitude of the land-only $<s(k)>$ time series for the MSLP anomalies is only marginally less than that of the corresponding global time series.

From Figure 10, it is evident that the spatial reproducibility of the seasonal precipitation anomalies fluctuates markedly in time. The largest peaks in $<s(k)>$ coincide with occurrences of El Niños in 1982/83 and 1986/87; in some cases, smaller peaks occur at times of La Niñas in 1984/85 (Figures 10a and 10c) and in the latter part of 1988 (Figure 10a). At such times, the reproducibility of the continental anomalies is more pronounced in the Tropics (Figure 10c), but a substantial extratropical signal associated with the exceptionally intense 1982/83 El Niño is present in MAM 1983 (Figures

10b). Most of the other model surface anomalies also display qualitatively similar increases in extratropical spatial reproducibility during MAM 1983 (not shown).

A propensity for global predictability to peak in the MAM season following onset of an El Niño in the latter part of the previous year has been noted by many other investigators (e.g., Brankovic et al. 1994, Anderson and Stern 1996, Bengtsson et al. 1996, Zwiers 1996). The timing is thought to be related to the maximal extent of tropical SSTs exceeding 28 degrees Celsius (a threshold for convection), and a dynamical basic state that expedites tropical-extratropical interactions (Hoerling et al. 1995, Higgins and Halpert 1997, Sud et al. 1999).

Figure 11, for example, displays the MAM 1983 field of the mean anomaly $\overline{A}$ (averaged over 6 model realizations) of MSLP, surface air temperature, and precipitation. Shaded areas indicate where these anomalies are significantly different (at the 95% confidence level) from other MAM seasons in years without El Niños--that is, in all other years of the simulation decade except 1987. (The significance test employs a nonparameteric Kolmogorov-Smirnov statistic that is well-suited for non-Gaussian variables such as precipitation--cf. Anderson and Stern 1996.)

Several significant anomalies commonly associated with El Niños (e.g. Quiroz 1983, Ropelewski and Halpert 1987, Barnett 1988, Kiladis and Diaz 1989, Kane 1997) are evident in Figure 11. These include 1) pronounced Southern Oscillation and North Pacific surface pressure anomalies (Figure 11a), 2) anomalous warming of tropical continents and of northern/western North America, with cooling of boreal continental interiors (Figure 11b), and 3) abnormally wet/dry conditions in the eastern/western tropical Pacific (Figure 11c). (The canonical precipitation anomaly centered in the southeastern U.S. during El Niños is more evident in the analogous map for the preceding DJF season, not shown here.) The temperature anomalies (Figure 11b) are most widespread over the extratropical continents, and the precipitation anomalies (Figure 11c) are least so, but the latter are more extensive

during this El Niño case than is the model norm.

Figure 12 depicts the overall spatial reproducibility of the model land-surface anomalies, as indicated by aggregate ensemble statistics $<S>$ and ΔS (computed over N = 15 samples of s(k) for each anomaly), where S is the temporal mean of a sample time series of spatial cross-correlations s(k), for seasons k=1, 2, ..., 39. In Figure 12a, $<S>$ (black bars) and ΔS (gray bars) are determined from s(k) computed over all continental surfaces, while in Figure 12b, the s(k) pertain only to tropical (30S to 30N) land points.

In contrast to their relatively low aggregate temporal reproducibility (Figure 9a), anomalies of precipitation (**pr**), evaporation (**evs**), and sensible heat flux (**hfss**) exhibit relatively high aggregate spatial reproducibility (Figure 12a). Conversely, all-land anomalies of MSLP **(psl)** , ground temperature (**tg**) and surface air temperature (**tas**) show relatively low overall spatial reproducibility even though their temporal reproducibility is comparatively high (Figure 9a). This apparent spatio-temporal asymmetry *may* be physically based (e.g. indicating that the continental **pr**, **evs**, and **hfss** land anomaly patterns are more spatially reproducible than those of **psl**, **tas**, and **tg**), or at least it may be representative of present-day GCM land-surface simulations (e.g. Phillips et al. 2000). Instead, however, this asymmetry could mostly reflect differences in statistical sampling and/or weighted averaging, in that R is an area-weighted spatial average over several thousand grid points, while S is a simple temporal average over just 39 unweighted seasonal pattern correlations. Because R thus is likely to be more a more robust statistic than S, the summary information in Figure 9 probably should be emphasized more than that of Figure 12 for purposes of ranking the overall reproducibility of the model's land-surface anomalies. In any event, there is not much difference in the aggregate

$<S>$ and ΔS values among the surface anomalies computed only over tropical land points (Figure 12b), indicating again that it is their extratropical patterns that are especially chaotic.

5. Correspondence to Conventional Measures of Potential Predictability

The ensemble-mean cross-correlations $<r(i,j)>$ and $<s(k)>$ are likely to be comparatively robust measures of reproducibility since, for the given n = 6 model realizations, these metrics are calculated from N = 15 samples. However, *r* and *s* themselves are computed from gridded seasonal data that are not statistically independent, since they are serially/spatially correlated with data at neighboring seasons/grid points. This complicates estimation of threshold values of *r* or *s* that would signify likely instances of PP (at, say, a 95% confidence level). Although such thresholds could be empirically determined via resampling techniques such as those described by Livezey and Chen (1983), the resulting criteria would be variable-specific (e.g. Phillips 1987, 1992). Instead, following a more heuristic approach, it will be shown that inferences of PP based on *r* and *s* compare well with conventional measures of predictability that have more clearcut likelihood thresholds.

Two such conventional predictability measures are considered here: a fractional variance statistic $f(i,j)$ and a forecast skill score $c(k)$ derived from the spatial pattern correlations of simulated vs. observed anomalies. Note, however, that both these metrics are obtained at a cost of reducing the effective sample size from N = 15 to n = 6 .

a. Fractional variance measure

The fractional variance estimate of PP (cf. Madden 1981) is computed from the fraction *f* of the total variability σ_T^2 (i,j) of climate variable V at each grid point (i,j) that is attributable to the prescribed ocean boundary forcing $\sigma_B^2(i,j)$, as opposed to the internal variability $\sigma_I^2(i,j)$ that is

assumed to be unpredictable :

$$f(i,j) = \sigma_B^2(i,j) / \sigma_T^2(i,j) \text{ , where } \sigma_T^2(i,j) = \sigma_B^2(i,j) + \sigma_I^2(i,j)$$

In this framework, $\sigma_B^2(i,j) > \sigma_I^2(i,j)$ (or $f(i,j) > 0.5$) is a conservative criterion for implying likely instances of PP. Analogous to the temporal reproducibility cross-correlation $\langle r \rangle$, the statistic f is computed at each grid point (i,j); in contrast to $\langle r(i,j) \rangle$, however, $f(i,j)$ usually is determined as a function of climatological season (i.e. average MAM, JJA, SON, and DJF), thereby permitting investigation of possible seasonal variations in PP.

Following Kumar and Hoerling (1995), $\sigma_B^2(i,j)$ is estimated from the ensemble-mean interannual variability of V in climatological season τ, and $\sigma_I^2(i,j)$ from the intraensemble variability of the different model realizations of τ, where in both cases the sample size is n = 6. Maps of $f(i,j)$ indicate that the PP of the model's MSLP displays greater extratropical extent than that of any other land-surface variable, notably over western North America in JJA and over Australia in DJF (Figure 13). The spatial pattern of PP estimated from metric $f(i,j)$ is very similar to what is implied by the temporal reproducibility statistic $\langle r(i,j) \rangle$ (Figure 8a). The surface air temperature also exhibits PP over parts of North America in climatological MAM (not shown), a pattern reminiscent of that reported by Zwiers (1996) for the 850 hPa temperature field in another GCM. More typical of the land-surface variables, however, are the maps of $f(i,j)$ for precipitation (not shown), which indicate very few instances of PP outside the deep Tropics.

b. Anomaly correlation measure

In operational seasonal forecasting, a common measure of predictive skill is the "anomaly correlation" $c(k)$ of a simulated variable with respect to its observation in season k. For each of the k = 39 seasons in the AMIP decade, $c(k)$ is computed from the spatial correlation of the ensemble-

mean model surface anomaly <A> (in this case, an arithmetic average over an ensemble of n = 6 members) against an available realization of the observed anomaly O :

$$c(k) = \sum_{i,j}^{I \times J} \{<A(i,j,k)> - <[A(k)]>\} * \{O(i,j,k) - [O(k)]\} / \{(IxJ) * <\sigma_{SA}(k)> * \sigma_{SO}(k)\}$$

Here <[A (k)]> and [O (k)] denote spatial averages of <A(i,j,k)> and O(i,j,k) over the relevant domain (e.g. global, all-land, etc.) in each season k, while $<\sigma_{SA}(k)>$ and $\sigma_{SO}(k)$ are the corresponding spatial standard deviations.

Estimates of O(i,j,k) for the chosen model surface variable of Table 1 were all obtained from the NCEP-NCAR reanalysis data, in order to permit an independent but consistent comparison of c(k) values across the model variables. Following Zwiers (1996), the uncertainty δc(k) in the anomaly correlation c(k) also was estimated as the average scatter among the spatial pattern correlations c^m of each realization m = 1,2, ..., n = 6 of the model anomaly A^m(i,j,k) with the (single) NCEP-NCAR estimate of the observational anomaly O(i,j,k):

$$\delta c(k) = \sum_{m=1}^{n=6} \{c(k) - c^m(k)]\}^2 / n\}^{1/2}$$

where

$$c^m(k) = \sum_{i,j}^{I \times J} \{A^m(i,j,k) - [A^m(k)]\} * \{O(i,j,k) - [O(k)]\} / \{(IxJ) * \sigma^m_{SA}(k) * \sigma_{SO}(k)\}$$

Time series of c(k) and its uncertainty for the model's MSLP (**psl**) are displayed in Figure 14. The threshold value c = 0.50 (denoted by dot-dash lines) indicates a minimum level of effective forecast skill, in the sense that the prediction may be of practical value (Kumar and Hoerling 1995).

According to this criterion, effectively skillful seasonal forecasts of both global (Figure 14a) and all-land MSLP (Figure 14b) occur in only a few seasons during the AMIP decade (e.g. coincident with the 1982/83 El Niño and the 1988/89 La Niña). The prediction of continental pressure is more frequently skillful in the Tropics, sometimes for several seasons in a row (Figure 14c), but even here there are wide swings in *c* with season k, and a number of negative-valued correlations. (The *c* statistic is known to be quite "unforgiving", often yielding negative values--cf. Zwiers 1996.) In contrast, the anomaly correlations for precipitation (not shown) are positive-valued for all seasons k, but there is only one instance of an effectively skillful global forecast, and there are no skillful predictions of continental precipitation, even in the Tropics.

The distinction between the more frequently skillful but "erratic" pressure correlations and the generally less skillful but "consistent" precipitation correlations is captured by the all-land temporal means *C* and associated uncertainties, which are displayed in Figure 15a along with those of the other nine model surface anomalies. Here the all-land MSLP (**psl**) anomaly exhibits the lowest mean *C*, and its uncertainty is also comparatively large, while the all-land *C* for the precipitation (**pr**) anomaly is the highest of any in the model, and its uncertainty is also relatively low. These results are qualitatively similar to the all-land summary statistics of spatial reproducibility of the surface anomalies in Figure 12a, where the mean $\langle S \rangle$ of the time series $s(k)$ for precipitation and MSLP bracket the range of those for the other surface anomalies. It is also noteworthy that the relative ranking of the all-land anomalies in Figure 15a (in descending order of all-land *C* values) roughly corresponds to what was obtained from the all-land $\langle S \rangle$ values in Figure 12a.

When evaluation is limited to the tropical continents (Figure 15b), *C* for every surface anomaly is seen to increase from its all-land value (compare Figures 15a and 15b). The increase is especially

dramatic for MSLP (**psl**), but is also substantial for ground and surface air temperatures (**tg** and **tas**) and for the surface wind stresses (**tauu** and **tauv**). The relative reductions of the summary statistics of certain model tropical land anomalies (e.g. soil moisture **mrso** and turbulent fluxes **evs** and **hfss**) that occur in passing from Figure 12b to Figure 15b highlight the distinctions between predictability metrics that are model-centric (e.g. *s* and *S*) and more demanding measures (e.g. *c* and *C*) that are based on reference data external to the model. However, to attribute all these apparent reductions in predictability metrics to the shortcomings of the ECMWF model seems unjustified, especially for land-surface processes such as turbulent fluxes and soil moisture that are mainly determined by the parameterizations of the NCEP analysis model, or that may exhibit biases introduced by the data assimilation techniques (Kanamitsu et al. 2002). It should be emphasized as well that evaluation of model predictability using observational data is an inherently probabilistic exercise, since nature only provides a single realization of the observed climate that need not correspond to the ensemble-mean seasonal forecast of even the hypothetically "perfect" model.

Thus, all measures of model predictability considered here display different strengths and weaknesses. The reproducibility metrics r and s maximally utilize the information contained in the available realizations by considering all N = n!/2!(n-2)! independent correlations among the n ensemble members; however, statistical confidence levels are model- and variable-specific, and therefore are not determinable *a priori*. The confidence levels are well-defined when estimating PP by means of a fractional variance statistic *f*, but at a cost of limiting the effective number of samples to the ensemble size n. The anomaly correlation *c* provides a stringent test of a model's actual predictive capabilities, but this metric may be subject to substantial observational uncertainties and, like statistic *f*, draws on a more limited sample population. Existing alternative procedures for inferring PP that

attempt, for example, to identify modes of climate variability associated with slowly varying boundary conditions (Fredericksen et al. 2003, Zheng and Fredericksen 2004) also merit further investigation. Such alternatives are also likely to have their own advantages and disadvantages, however.

6. Discussion

This study analyzes the initial-condition sensitivity of eleven seasonal surface variables in an ensemble of six ECMWF (cycle 36) model simulations of decadal climate, with common specification of AMIP radiative and ocean boundary conditions. To quantify this sensitivity, spatio-temporal measures of reproducibility of these surface variables are computed, with particular attention to their values on land. The reproducibility of the surface variables when the seasonal cycle is included or excluded also is separately considered, since the former configuration is relevant for climate model intercomparison, and the latter for seasonal climate prediction. Possible ramifications of this study for both these endeavors are next discussed.

a. Implications for model intercomparisons

Because chaos complicates the analysis of a climate simulation, it is well to focus a model intercomparison project on variables that display low sensitivity to initial conditions. To the extent that the ECMWF cycle 36 model is representative, the present study implies that this condition will be met by most land-surface variables that include the seasonal cycle, with the possible exception of precipitation and surface wind stresses. However, the decadal seasonal climatologies of even these highly chaotic variables proved to be spatially reproducible to a high degree. This suggests that the comparison of seasonal land-surface climatologies across different atmospheric GCMs is potentially a meaningful exercise, provided that common radiative and ocean boundary conditions are specified, such as in the AMIP experiments. (The intercomparison of model *monthly* climatologies seems somewhat more

problematical, however, in view of their greater sensitivity to initial conditions.)

A more fundamental problem for intercomparison projects focussing on the effects of different coupled land-surface schemes is that simulations of continental climate reflect the biases of the atmospheric model surface forcings (e.g. net radiation or precipitation), thus making it difficult to assess the degree to which the different embedded land-surface schemes influence the simulation. In principle, the impact of these biases could be reduced by intercomparing the models' land-surface anomalies (i.e. the departures from the respective seasonal cycles). However, if only single realizations of each model's climate were available, the present study implies that such a strategy would prove inadequate, given the highly chaotic character of such anomalies outside the deep Tropics.

Such an approach would only be effective if multiple realizations of seasonal land-surface climate were available for each model, so that the reproducible part of each anomaly could be estimated from its ensemble mean; however, many realizations (probably more than 6, as implemented here) would be needed for an accurate estimate (Barnett 1995, Wehner 2000). Because such an ambitious ensemble intercomparison study is still difficult to implement, it may be advisable to explore other methods for *a posteriori* reduction of the effects of model biases. For example, one promising approach after Gedney et al. (2000) is to construct dimensionless ratios of response/ forcing variables to elucidate essentially different behaviors of coupled land-surface schemes.

Yet another fundamental challenge to analyzing the effects of different land-surface schemes in coupled mode is that the strength and spatial pattern of the land-atmosphere coupling is likely to be model-dependent (e.g. Koster et al. 2002, 2004). This implies that the results of the present study depend to unknown degree on the particular biases and land-atmosphere coupling characteristics of the ECMWF cycle 36 model. The way forward in resolving such ambiguities therefore probably

will entail careful analysis of additional well-designed ensemble experiments with a diverse collection of coupled land-atmosphere models.

b. Implications for seasonal climate prediction

In the present study the potential predictability (PP) of the ECMWF cycle 36 model's land-surface climate is inferred by computing measures of the spatio-temporal reproducibility of the seasonal anomalies of like variables in the ensemble. A considerable range in the spatial extent of the implied PP of the model's land-surface processes is seen, with continental surface pressure and temperature generally exhibiting wider-ranging PP than the hydrological variables or the turbulent fluxes. Surface pressure and temperature also manifest PP over the extratropical continents, albeit only in selected regions and seasons. Moreover, although the overall PP of the model's seasonal continental climate is substantial only in the deep Tropics, the spatial reproducibility of the simulated extratropical seasonal continental climate is seen to be perceptibly enhanced by ENSO events, notably during the intense 1982/83 El Niño. These results are generally consistent with those implied by more conventional measures of PP.

These rather sobering implications for the prospects of predicting extratropical continental climate from "perfect" knowledge of SSTs may be tempered somewhat by consideration of possible means for enhancing forecast skill during a particular season. First, in contrast to this study's numerical experiments, operational seasonal predictions usually proceed from initial conditions specified from the observed atmospheric state (e.g. as provided by analysis products). Hence, some additional predictive skill is often imparted by atmospheric persistence, at least in the earlier stages of the seasonal forecast (e.g. Dirmeyer et al. 2001).

There is also considerable evidence (e.g. Huang et al. 1996; Wang and Kumar 1998; Cohen and

Entekhabi 1998; Ferranti et al. 1999; Dirmeyer 2001; Koster et al. 2000, 2002, 2004) that seasonal predictive skill can be enhanced by specification of other sources of climate memory such as snow cover and near-surface soil moisture (where the latter is treated as a predictor variable rather than a predictand, as in the present study). Additional knowledge of the ocean state (other than that provided by SSTs) also may yield an enhanced set of seasonal climate predictors (e.g. Phillips 1992, Griffies and Bryan 1997). Further advances on these fronts no doubt will be closely tied to integrated remote sensing and modeling initiatives promoted by the World Climate Research Programme (WCRP) and similar collaborations.

Seasonal forecasts also may be enhanced by intelligent application of statistical forecasting techniques: because the land-surface processes are correlated, skillful dynamical forecasts of only some variables can impart statistical knowledge of others (Barnston and Smith 1996). Enhanced predictive skill may also be realized by reducing systematic model errors through statistical correction techniques (Smith and Livezey 1999, Feddersen et al. 1999), by combining dynamical and statistical forecasting techniques (Anderson et al. 1999), or by utilizing ensemble forecasts from more than one model (Mason et al. 1999). For example, Koster et al. (2004) report "consensus" results from 12 coupled land-atmosphere models which indicate that during boreal summer consistently strong coupling between soil moisture and precipitation occurs over India, Sub-Saharan and Equatorial Africa, and the Great Plains of North America.

Finally, it should be reiterated that the potential predictability of land-surface climate also depends partly on model performance, which is a complex function of the simulated land-atmosphere interactions. Hence, with future advances in computer technology facilitating increasingly realistic simulations of continental processes, there is reason to anticipate further progress in our ability to predict seasonal land-surface climate.

Acknowledgments

Special thanks are due to Ben Santer for recommending the reproducibility metrics used in this study. The cooperation of the European Center for Medium-Range Weather Forecasts in making available their model also is gratefully acknowledged. This work was performed under the auspices of the U.S. Department of Energy's Office of Science, Biological and Environmental Reseach (BER) program by the University of California Lawrence Livermore National Laboratory under Contract W-7405-ENG-48.

References

Anderson, J.L., and W.F. Stern, 1996: Evaluating the potential predictive utility of ensemble forecasts. *J. Climate*, **9**, 260-269.

Anderson, J., H. van den Dool, A. Barnston, W. Chen, W. Stern, and J. Ploshay, 1999: Present-day capabilities of numerical and statistical models for atmospheric extratropical seasonal simulation and prediction. *Bull. Amer. Meteor. Soc*., **80**, 1349-1361.

Barnett, T.P., 1995: Monte Carlo climate forecasting. *J. Climate*, **8**, 1005-1022.

Barnett, T.P., 1988: Variations in near-global sea level pressure: Another view. *J. Climate*, **1**, 225-230.

Barnston, A.G., 1994: Linear statistical short-term climate predictive skill in the Northern Hemisphere. *J. Climate*, **7**, 1513-1564.

Barnston, A.G., and T.M. Smith, 1996: Specification and prediction of global surface temperatures and precipitation from global SST using CCA. *J. Climate*, **9**, 2660-2697.

Beljaars, A.C.M., P. Viterbo, and M.J. Miller, 1996: The anomalous rainfall over the United States during July 1993: Sensitivity to land surface parameterization and soil moisture anomalies. *Mon. Wea. Rev*., **124**, 362-383.

Bengtsson, L., and J. Shukla, 1988: Integration of space and in situ observations to study global climate change. *Bull. Amer. Meteor. Soc.*, **69**, 1130-1143.

Bengtsson, L., K. Arpe, E. Roeckner, U. Schulzweida, 1996: Climate predictability experiments with a general circulation model. *Climate Dyn.*, **12**, 261-278.

Betts, A.K., J.H. Ball, and A.C.M. Beljaars, 1993: Comparison between the land surface response of the ECMWF model and the FIFE-1987 data. *Quart. J. Roy. Meteor. Soc*., **119**, 975-1001.

Blondin, C., 1989: Research on land surface parameterisation schemes at ECMWF. In Proceedings of the Workshop on Parameterisation of Fluxes over Land Surface, European Centre for Medium-Range Weather Forecasts, Reading, England.

Brankovic, C., T.N. Palmer, and L. Ferranti, 1994: Predictability of seasonal atmospheric variations. *J. Climat*e, **7**, 217-237.

Carson, D.J. 1998: Seasonal forecasting. *Quart. J. Roy. Meteor. Soc*., **124**, 1-26.

Chen, W.Y., and H.M. van Van den Dool, 1997: Atmospheric predictability of seasonal, annual, and decadal climate means and the role of the ENSO cycle: A model study. *J. Climate*, **10**, 1236-1254.

Chervin, R. M., 1986: Interannual variability and seasonal climate predictability. *J. Atmos. Sci.,* **43**, 233-251.

Cohen, J., and D. Entekhabi, 1998: Eurasian snow cover variability and Norther Hemisphere climate predictability. *Geophys. Res. Letters*, **26**, 345-348.

Cubasch, U., B.D. Santer, A. Helibach, G. Hegerl, H. Höck, E. Maier-Reimer, U. Mikolajewicz, A. Stoessel, and R. Voss, 1994: Monte Carlo climate change forecasts with a global coupled ocean-atmosphere model. *Clim. Dynam.*, **10**, 1-19.

Delworth, T.L., and T.R. Knutson, 2000: Simulations of early 20th century global warming. *Science*, **287 (5461)**, 2246-2250.

Dirmeyer, P.A., 2001: An evaluation of the strength of land-atmosphere coupling. *J. Hydrometeor.*, **2**, 329-344.

Dirmeyer, P.A., M.J. Fennessy, and L. Marx, 2001: Near surface boreal summer climate as simulated by three general circulation models. COLA Technical Report #100, Center for Ocean-Land-Atmosphere Studies, Calverton, Maryland 20705-3106, USA, 36 pp.

Dix, M.R., and B.G. Hunt, 1995: Chaotic influences and the problem of deterministic seasonal predictions. *Int. J. Climatol.*, **15**, 729-752.

Ebisuzaki, W., 1995: The potential predictability in a 14-year GCM simulation. *J. Climate*, **8**, 2749-2761.

ECMWF Research Department, 1988: ECMWF forecast model, adiabatic part (2nd edition). European Centre for Medium-Range Weather Forecasts, Reading, England.

ECMWF Research Department, 1991: ECMWF forecast model, physical parameterisation (3rd edition). European Centre for Medium-Range Weather Forecasts, Reading, England.

Feddersen, H., A. Navarra, and M.N. Ward, 1999: Reduction of model systematic error by statistical correction for dynamical seasonal prediction. *J. Climate*, **12**, 1974-1989.

Ferranti, L. J.M. Slingo, T.N. Palmer, and B.J. Hoskins, 1999: The effect of land-surface feedbacks on the monsoon circulation. *Quar. J. Roy. Meteor. Soc.*, **125**, 1527-1550.

Frederiksen, C.S., A.P. Kariko, and X. Zheng, 2003: A technique for extracting potentially predictably patterns from climate data. Anziam J., 44(E), pp. C160-C179. Accessible online at http://anziamj.austms.org.au/V44/CTAC2001/Fred .

Gates, W.L., 1992: AMIP: The Atmospheric Model Intercomparison Project. *Bull. Amer. Meteor. Soc.*, **73**, 1962-1970.

Gates, W.L.,J.S. Boyle, C. covey, C.g. Dease, C.M. Doutriaux, R.S. Drach, M. Fiorino, P.J. Gleckler, J.J. Hnilo, S.M. Marlais, T.J. Phillips, G.L. Potter, B.D. Santer, K.R. Sperber, K.E. Taylor, and D.N. Williams, 1999: An overview of the results of the Atmospheric Model Intercomparison Project (AMIP I). *Bull. Amer. Meteor. Soc.*, **80**, 29-55.

Gedney, N., P.M. Cox, H. Douville, J. Polcher, and P.J. Valdes, 2000: Characterising GCM land surface schemes to understand their responses to climate change. *J. Climate*, **13**, 3066-3079.

Gibson, J.K., P. Kallberg, S. Uppala, A. Noumura, A. Hernandez, and E. Serrano, 1997: ERA Description. ECMWF Re-Analysis Project Report Series, 1. European Centre for Medium-range Weather Forecasts, Reading, UK, 77 pp.

Griffies, S.M., and K. Bryan, 1997: A predictability study of simulated North Atlantic multidecadal variability. *Climate Dyn.*, **13**, 459-487.

Hansen, J., and coauthors, 1997: Forcings and chaos in interannual to decadal climate change. *J. Geophys. Res.*, **102**, 25679-25720.

Henderson-Sellers, A., K. McGuffie, and A.J. Pitman, 1996: The Project for Intercomparison of Land-Surface Parameterization Schemes: 1992 to 1995. *Climate Dyn*., **12**, 849-859.

Higgins, W. and M. Halpert, 1997: Oceanic and atmospheric anomalies in the global Tropics: The 1997 El Niño. In the Proceedings of the Twenty-Second Annual Climate Diagnostics and Prediction Workshop, Berkeley, California, 6-10 October 1997, pp. 1-5.

Hoerling, M.P., M. Ting, and A. Kumar, 1995: Zonal-flow stationary wave relationship during El Niño: Implications for seasonal forecasting. *J. Climate*, **8**, 1838-1852.

Huang, J., H.M. van den Dool, and K.P. Georgakakos, 1996: Analysis of model-calculated soil moisture over the United States (1931-1993) and applications to long range temperature forecasts. *J. Climate*, **9**, 1350-1362.

Huffman, G.H. and D.T. Bolvin, 2004: GPCP version 2 combined precipitation data set documentation. Accessible online at http://www1.ncdc.noaa.gov/pub/data/gpcp/v2/documentation/V2_doc .

Kalnay, E. and coauthors, 1996: The NMC/NCAR 40-year reanalysis project. *Bull. Amer. Meteor. Soc.*, **77**, 437-471.

Kanamitsu,M., W. Ebisuzaki, J. Woollen, S-K. Yang, J.J. Hnilo, M. Fiorino, and G.L. Potte, 2002: NCEP-DOE AMIP-II Reanalysis (R-2). *Bull. Amer. Meteor. Soc*., **83**, 1631-1643.

Kane, R.P., 1997: Relationship of El Niño-Southern Oscillation and Pacific sea surface temperature with rainfall in various regions of the globe. *Mon. Wea. Rev*., **125**, 1792-1800.

Kiladis, G.N., and H.F. Diaz, 1989: Global climatic anomalies associated with extremes in the Southern Oscillation. *J. Climate*, **2**, 1069-1090.

Koster, R.D., and M.J. Suarez, 1994: The components of a 'SVAT' scheme and their effects on a GCM's hydrological cycle. *Adv. Water Res*., **17**, 61-78.

Koster, R.D., M.J. Suarez, and M Heiser, 2000: Variance and predictability of precipitation at seasonal-to-interannual timescales. *J. Hydrometeor.* , **1**, 26-46.

Koster, R.D., P.A. Dirmeyer, A.N. Hahmann, R. Ijpelaar, L. Tyahla, P. Cox, and M.J. Suarez, 2002: Comparing the degree of land-atmosphere interactionin four atmospheric general circulation models. *J. Hydrometeorology*, **3**, 363-375.

Koster, R.D., P.A. Dirmeyer, Z. Guo, G. Bonan, E. Chan, P. Cox, C.T. Gordon, S. Kanae, E. Kowalczyk, D. Lawrence, P. Liu, C-H. Lu, S. Malyshev, B. McAvaney, K. Mitchell, D. Mocko, T. Oki, K. Oleson, A. Pitman, Y.C. Sud, C.M. Taylor, D. Verseghy, R. Vasic, Y. Xue, and T. Yamada, 2004: Regions of strong coupling between soil moisture and precipitation. Science, 305, 20 August 2004, 1138-1140.

Kumar, A., and M.P. Hoerling, 1995: Prospects and limitations of atmospheric GCM climate predictions. *Bull. Amer. Meteor. Soc.*, **76**, 335-345.

Kumar, A., M. Hoerling, M. Ji, A. Leetmaa, and P. Sardeshmukh, 1996: Assessing a GCM's suitability for making seasonal predictions. *J. Climate*, **9**, 240-247.

Leith, C.E., 1973: The standard error of time-averaged estimates of climatic means. *J. Appl. Meteor*., **12**, 1066-1069.

Liang, X-Z., K.R. Sperber, W-C. Wang, and A.N. Samel, 1997: Predictability of SST forced climate signals in two atmospheric general circulation models. *Climate Dyn., **13**, 391-415*.

Livezey, R.E., and W.Y. Chen, 1983: Statistical field significance and its determination by Monte Carlo techniques. *Mon. Wea. Rev*., **111**, 46-59.

Lorenz, E.N., 1964: The problem of deducing the climate from the governing equations. *Tellus*, **16**, 1-11.

Louis, J.-F., 1979: A parametric model of vertical eddy fluxes in the atmosphere. *Bound. Layer Meteor*., **17**, 187-202.

Madden, R., A., 1981: A quantitative approach to long-range prediction. *J. Geophys. Res*., **86**, 9817-9825.

Mason, S.J., L. Goddard, N.E. Graham, E. Yulaeva, L. Sun, and P.A. Arkin, 1999: The IRI seasonal climate prediction system and the 1997/98 El Niño event. *Bull. Amer. Meteor. Soc*., **80**, 1853-1873.

Maynard, K., and J. Polcher, 2003: Impact of land-surface processes on the interannual variability of tropical climate in the LMD GCM. *Climate Dyn*., **20**, 613-633.

Miller, M.J., T.N. Palmer, and R. Swinbank R, 1989: Parameterization and influence of subgrid-scale orography in general circulation and numerical weather prediction models. *Meteor. Atmos. Phys*., **40**, 84-109

Miller, M.J., A.C.M. Beljaars, and T.N. Palmer, 1992: The sensitivity of the ECMWF model to the parameterization of evaporation from the tropical oceans. *J. Climate*, **5**, 418-434.

Morcrette, J-J. , 1991: Radiation and cloud radiative properties in the ECMWF operational weather forecast model. *J. Geophys. Res.*, **96**, 9121-9132.

New, M., D. Lister, M. Hulme, and I. Makin, 2002: A high-resolution data set of surface climate over global land areas. *Climate Res*., **21**, 1-25.

Palmer, T.N., and D.L.T. Anderson, 1994: The prospects for seasonal forcasting--A review paper. *Quart. J. Roy. Meteor. Soc*., **120**, 755-792.

Pan, H.-L., and W-S. Wu, 1994: Implementing a mass flux convective parameterization package for the NMC medium-range forecast model. Preprints, *10th Conf. on Numerical Weather Prediction*, Portland, OR, Amer. Meteor. Soc., 96-98.

Phillips, T.J., 1987: The impact of sampling errors on the perceived response of a climate model to a seas surface temperature anomaly. *Atmos.-Ocean*, **25**, 177-196.

Phillips, T.J., 1992: An application of a simple coupled ocean-atmosphere model to the study of seasonal climate prediction. *J. Climate*, **5**, 1078-1096.

Phillips, T.J., 1994: A summary documentation of the AMIP models. PCMDI Report No. 18, Program for Climate Model Diagnosis and Intercomparison, Lawrence Livermore National Laboratory, Livermore, CA, 346 pp. Also accessible online at http://www-pcmdi.llnl.gov/modeldoc/amip1/ .

Phillips, T.J., A. Henderson-Sellers, P. Irannejad, K. McGuffie, H. Zhang, and the AMIP I Modeling Groups, 2000: Validation of land-surface processes in AMIP models: A pilot study. PCMDI Report No. 63, Program for Climate Model Diagnosis and Intercomparison, Lawrence Livermore National Laboratory, Livermore, CA, 20 pp. Also accessible online at http://www-pcmdi.llnl.gov/pcmdi/pubs/pdf/63.pdf .

Polcher, J. (ed.), 2000: Proceedings of the GEWEX/INSU International Workshop on Modelling Land-surface Atmosphere Interactions and Climate Variability, 4-8 October 1999, Gif-sur-Yvette, France, 134 pp.

Quiroz, R.S., 1983: The climate of the "El Niño" winter of 1982-83--A season of extraordinary climatic anomalies. *Mon. Wea. Rev.*, **111**, 1685-1706.

Ropelewski, C.F., and M.S. Halpert, 1987: Global and regional scale precipitation patterns associated with the El Niño/Southern Oscillation. *Mon. Wea. Rev*., **115**, 1606-1626.

Santer, B.D., 1988: Regional validation of general circulation models. *Clim. Res. Unit Publ*., **9**, 375 pp. Univ. of East Anglia, Norwich, England.

Scott, R., D. Entekhabi, R. Koster, and M. Suarez, 1997: Timescales of landsurface evapotranspiration response. *J. Climate,* 10, 559-566.

Shukla, J., J. Anderson, D. Baumhefner, C. Brankovic, Y. Chang, E. Kalnay, L. Marx, T. Palmer, D. Paolino, J. Ploshay, S. Schubert, D. Straus, M. Suarez, and J. Tribbia, 2000: Dynamical seasonal prediction. *Bull. Amer. Meteor. Soc.*, **81**, 2593-2606.

Slingo, J., 1987: The development and verification of a cloud prediction model for the ECMWF model. *Quart. J. Roy. Meteor. Soc*., **113**, 899-927.

Smith, T.M., and R.E. Livezey, 1999: GCM systematic error correction and specification of the seasonal mean Pacific-North America region atmosphere from global SSTs. *J. Climate*, **12**, 273-305.

Sperber, K.R., and T.N. Palmer, 1996: Interannual Tropical rainfall variability in general circulation model simulations associated with the Atmospheric Model Intercomparison Project. *J. Climate*, **9**, 2727-2750.

Sperber, K.R., and Participating AMIP Modelling Groups, 1999: Are revised models better models? A skill score assessment of regional interannual variability. *Geophys. Res. Letters*, **26**, 1267-1270.

Stern, W., and K. Miyakoda, 1995: The feasibility of seasonal forecasts inferred from multiple GCM simulations. *J. Climate*, **8,** 1071-1085.

Sud, Y.C., G.K. Walker, and K.-M. Lau, 1999: Mechanisms regulating sea-surface temperatures and deep convection in the Tropics. *Geophys. Res. Letters*, **26**, 1019-1022.

Tiedtke, M., 1989: A comprehensive mass flux scheme for cumulus parameterization in large-scale models. *Mon. Wea. Rev*., **117**,1779-1800.

Viterbo, P., and A.C.M. Beljaars, 1995: An improved land surface parameterization scheme in the ECMWF model and its validation. *J. Climate*, **8**, 2716-2748.

Wang, W., and A. Kumar, 1998: A GCM assessment of atmospheric seasonal predictability associated with soil moisture anomalies over North America. *J. Geophys. Res.*, **103**, 28,637-28,646.

Wang, X.L, and F.W. Zwiers, 1999: Interannual variability of precipitation in an ensemble of AMIP climate simulations conducted with the CCC GCM2. *J. Climate*, **12**, 1322-1335.

Wehner, M.F., 2000: Determination of the sampling size of AGCM ensemble simulations. *Climate Dyn.*, **16**, 321-331.

Yang, Z.-L., R.E. Dickinson, A. Henderson-Sellers, and A.J. Pitman, 1995: Preliminary study of spin-up processes in land-surface models with the first stage data of PILPS Phase 1(a). *J. Geophys. Res.*, **100**, 16553-16578.

Zheng, X., and C.S. Frederiksen, 2004: Variability of seasonal-mean fields arising from intraseasonal variability: part 1, methodology. *Climate Dyn*., **23**, 177-191.

Zwiers, F.W., 1987: A potential predictability study conducted with an atmospheric general circulation model. *Mon. Wea. Rev*., **115**, 2957-2974.

Zwiers, F.W., 1996: Interannual variability and predictability in an ensemble of AMIP climate simulations conducted with the CCC GCM2. *Climate Dyn.,* **12**, 825-847.

Zwiers, F.W., and V. Kharin, 1998: Intercomparison of interannual variability and potential predictability: an AMIP diagnostic subproject. *Climate Dyn*., **14**, 417-528.

Table 1: Surface variables considered in this study, listed in alphabetical order of their AMIP acronym, with units.

Surface Variable	**AMIP Acronym**	**Units**
Evaporation	**evs**	mm day^{-1}
Sensible Heat Flux	**hfss**	W m^{-2}
Soil Moisture	**mrso**	cm
Precipitation	**pr**	mm day^{-1}
Mean Sea-level Pressure (MSLP)	**psl**	hPa
Net Longwave Radiation	**rls**	W m^{-2}
Net Shortwave Radiation	**rss**	W m^{-2}
Surface Air Temperature*	**tas**	deg C
Eastward Wind Stress	**tauu**	Nt m^{-2}
Northward Wind Stress	**tauv**	Nt m^{-2}
Ground Temperature**	**tg**	deg C

* Temperature of the lowest vertical level of the model atmosphere, equivalent to a height ~ 30 meters above the surface.

** Equivalent to the surface skin temperature.

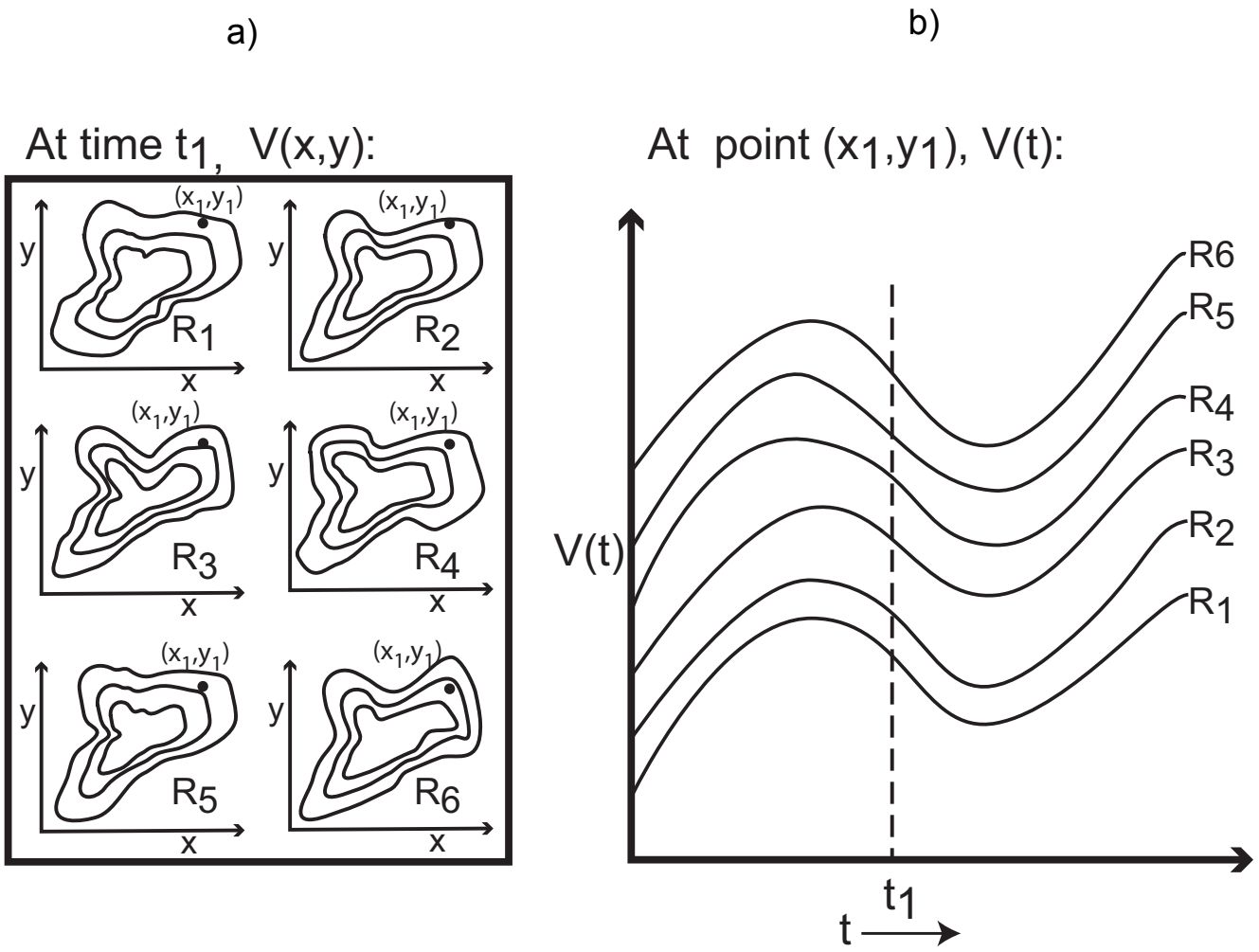


Figure 1: For surface variable V(x,y,t), high spatial reproducibility implies that at any time t_1 the pattern V(x,y) is similar across different statistically independent realizations R_1, R_2, ..., R_n of V(x,y,t) with somewhat different initial conditions, as shown in a). High temporal reproducibility inV(x,y,t) implies that the trajectories in time of V(t) of any point (x_1,y_1) in V(x,y,t) are similar across realizations R_1, R_2, ..., R_n , as shown in b).

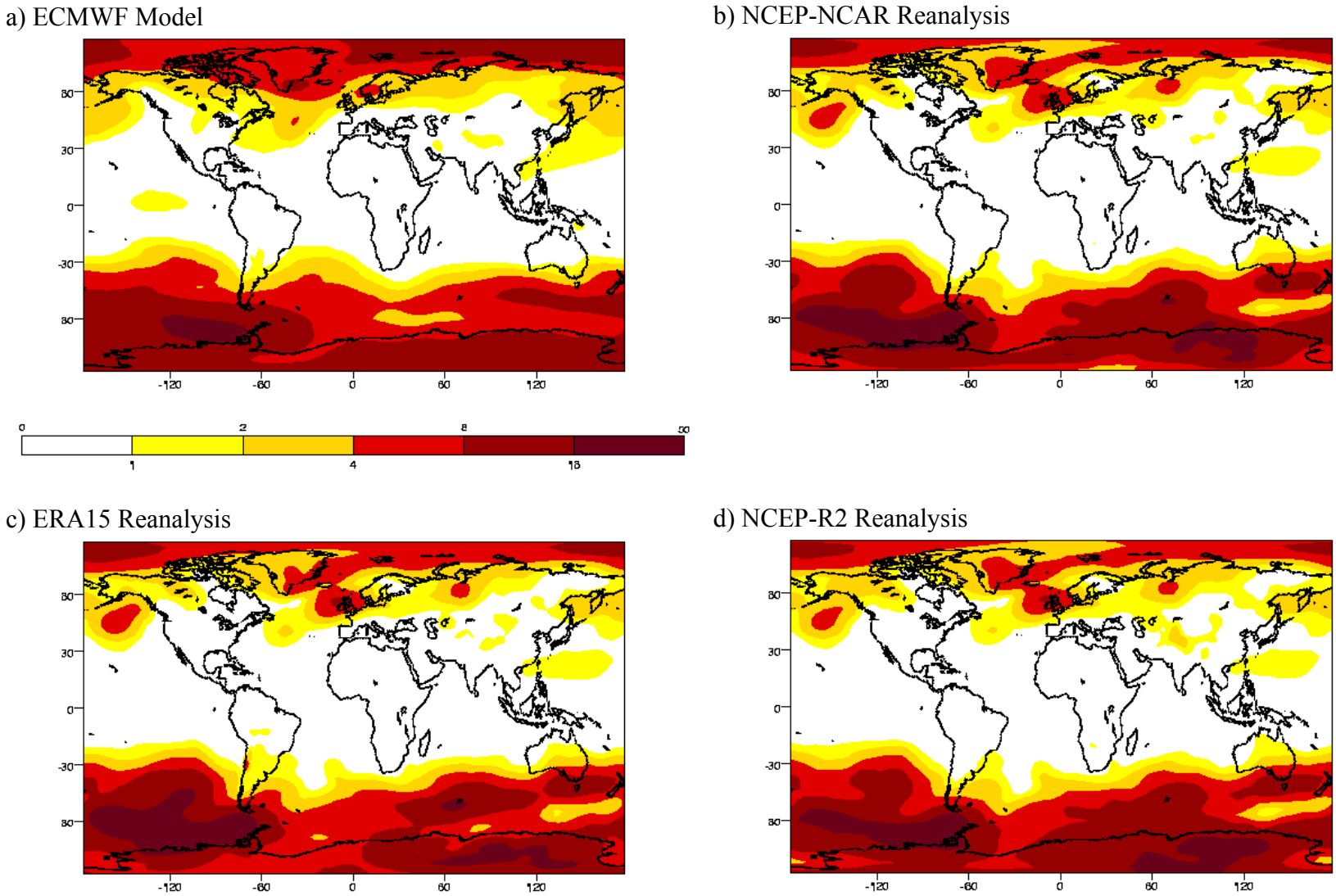


Figure 2: Ensemble-mean interannual variance of ECMWF cycle 36 model mean sea-level pressure (MSLP) in JJA (panel a) compared with that from the NCEP-NCAR reanalysis (panel b), the ERA15 reanalysis (panel c), and the NCEP-R2 reanalysis (panel d), all for the period 1979-1988. Here contour interval boundaries are 0.0, 1.0, 2.0, 4.0, 8.0, 16.0, and 50.0 $(hPa)^2$ and all validation data (panels b, c, and d) are interpolated to the spectral T42 Gaussian grid of the model (panel a).

Figure 2, continued

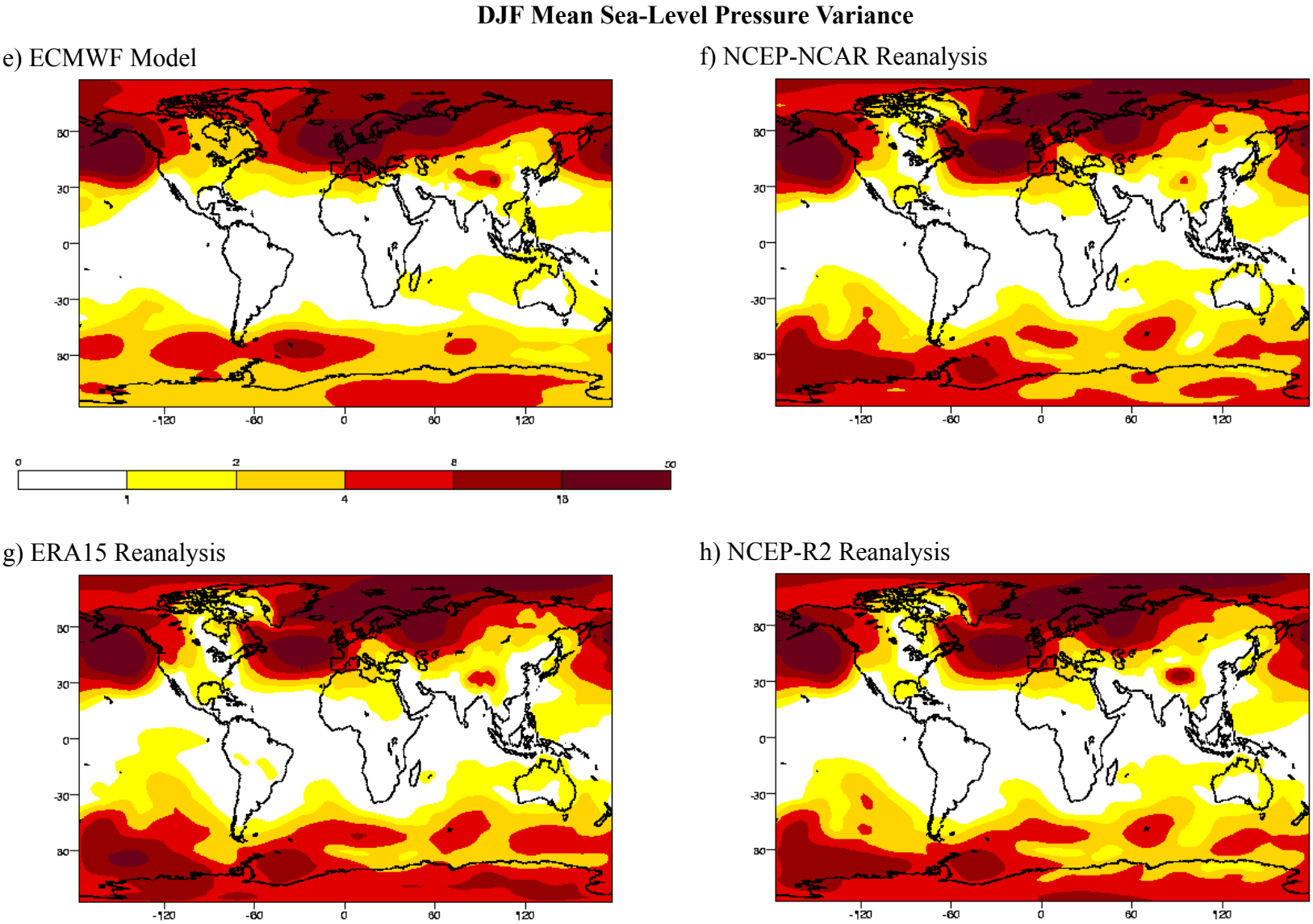


Figure 2, continued: Interannual variances of mean sea-level pressure (MSLP) as in panels a, b, c, and d, except for the DJF season.

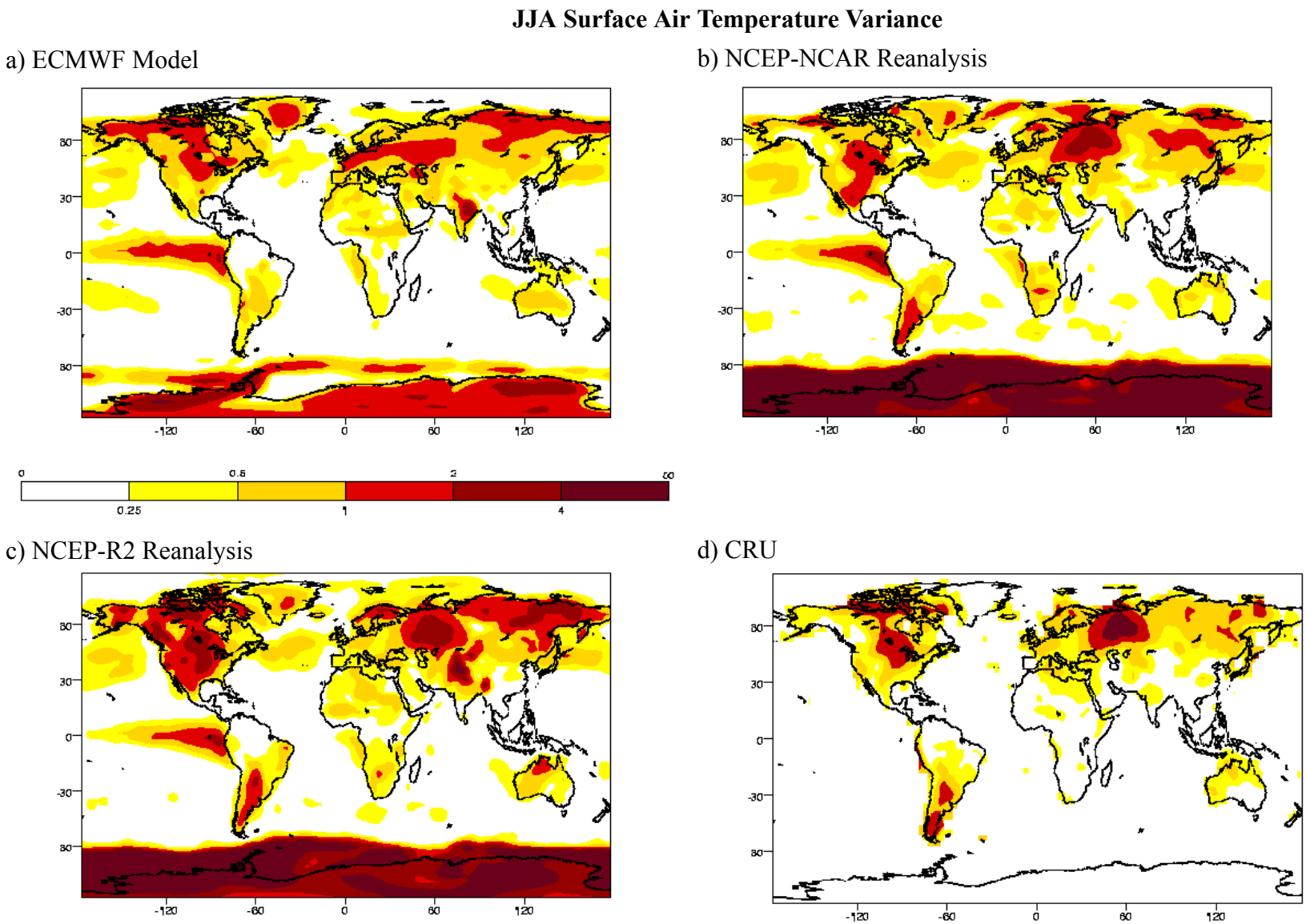


Figure 3: Ensemble-mean interannual variance of ECMWF cycle 36 model surface air temperature in JJA (panel a) compared with that from the NCEP-NCAR reanalysis (panel b), the NCEP-R2 reanalysis (panel c), and the CRU terrestrial data set (panel d), all for the period 1979-1988. Here contour interval boundaries are 0.0, 0.25, 0.5, 1.0, 2.0, 4.0, and 80.0 (mm day^{-1})^2 and all validation data (panels b, c, and d) are interpolated to ththe spectral T42 Gaussian grid of the model (panel a).

Figure 3, continued

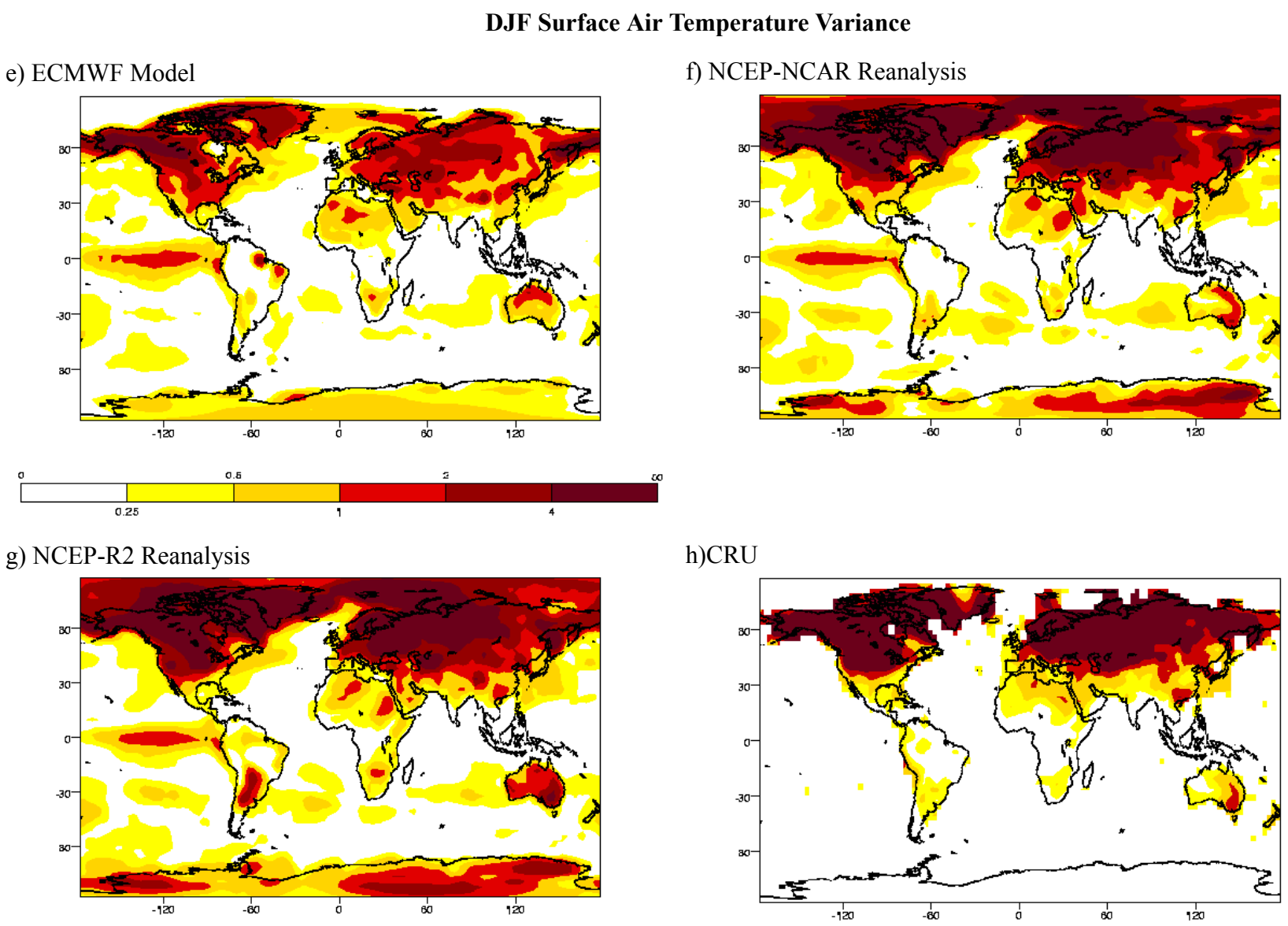
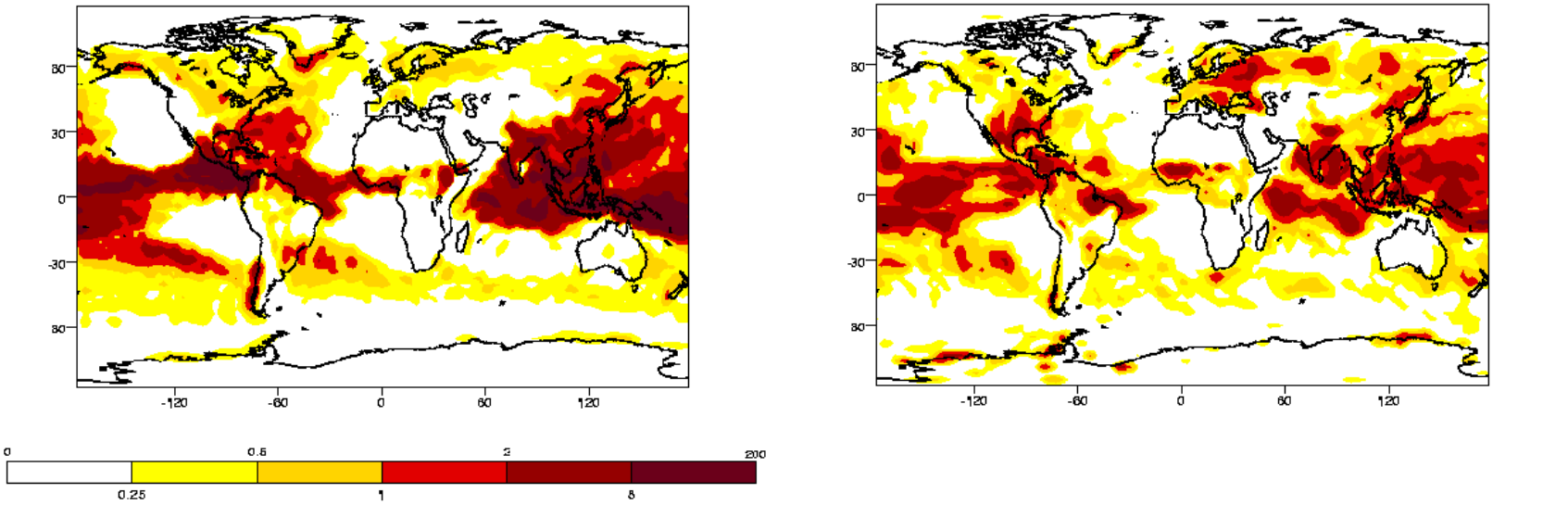


Figure 3, continued: Interannual variances of surface air temperature as in panels a, b, c, and d, except for the DJF season.

JJA Precipitation Variance

a) ECMWF Model

b) NCEP-NCAR Reanalysis



c) GPCP

d) CRU

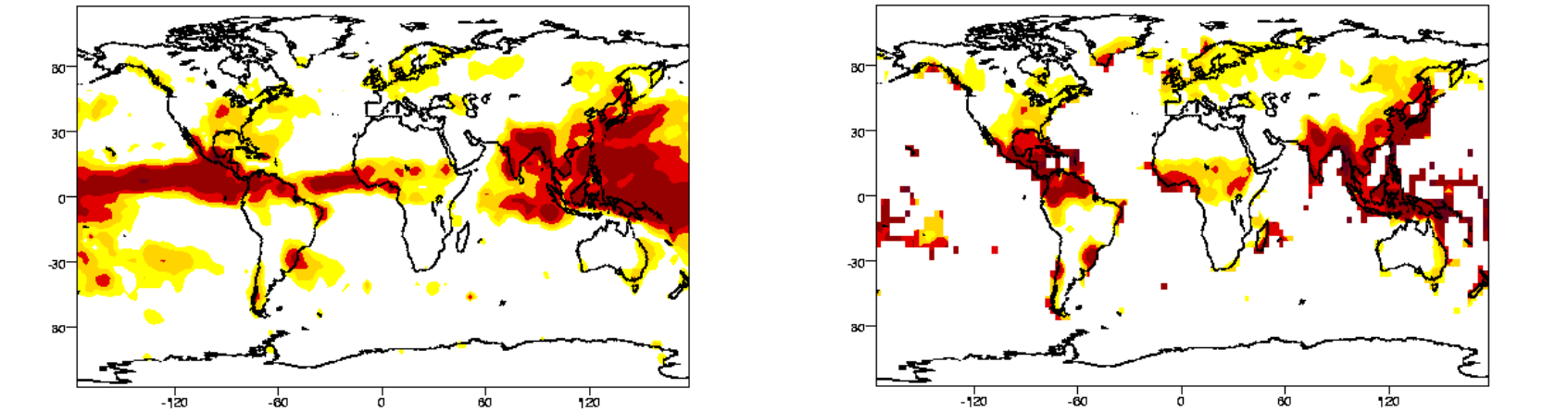


Figure 4: Ensemble-mean interannual variance of the ECMWF cycle 36 model precipitation in JJA (panel a) compared with that from the NCEP-NCAR reanalysis (panel b), the GPCP data set (panel c), and the CRU terrestrial (excluding Greenland and Antarctica) data set (panel d), all for the period 1979-1988. Here contour interval boundaries are 0.0, 0.25, 1.0, 2.0, 8.0, and 200.0 $(\text{mm day}^{-1})^2$ and all validation data (panels b, c, and d) are interpolated to the spectral T42 Gaussian grid of the model (panel a).

Figure 4, continued

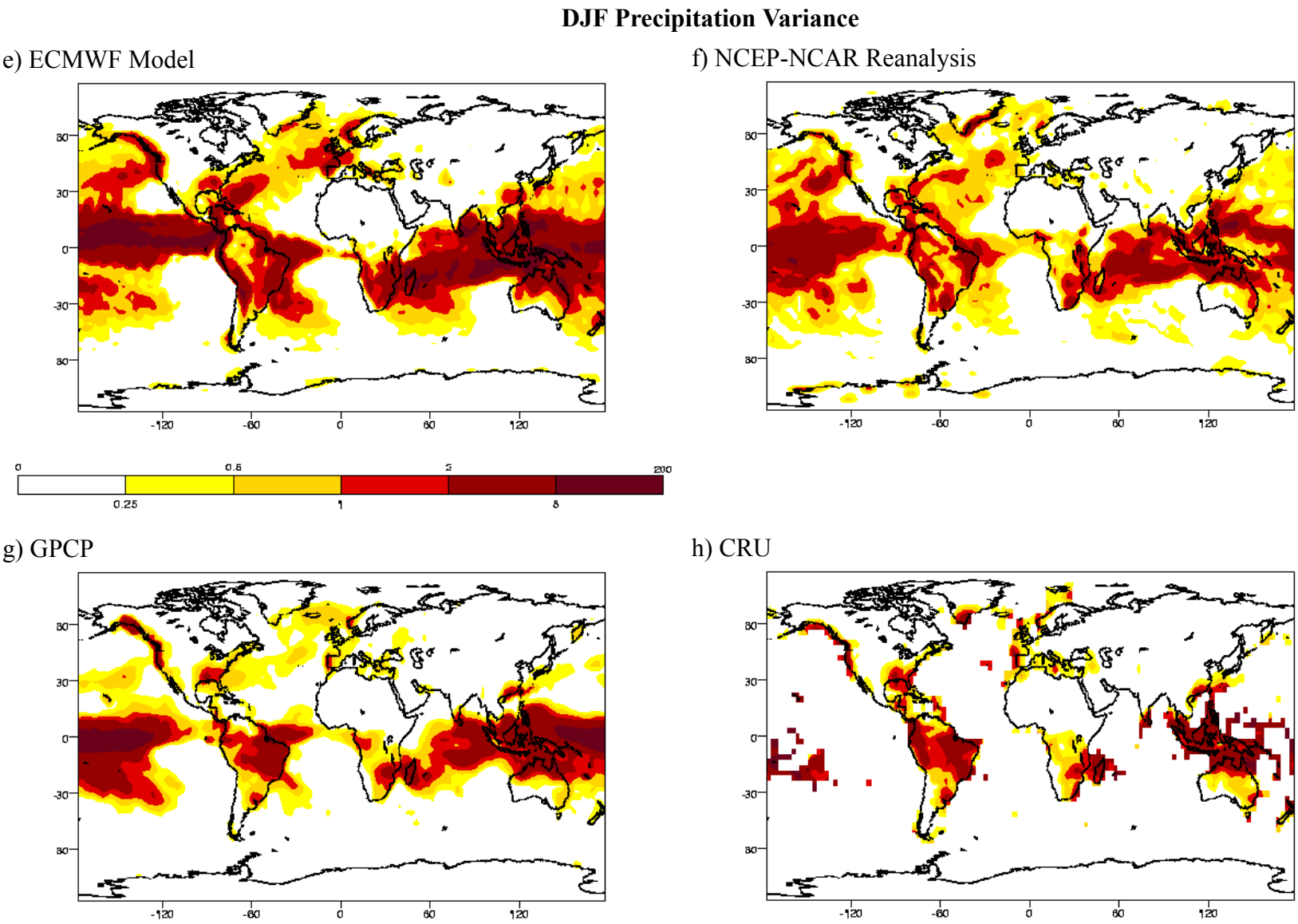


Figure 4, continued: Interannual variances of precipitation as in panels a, b, c, and d, except for the DJF season.

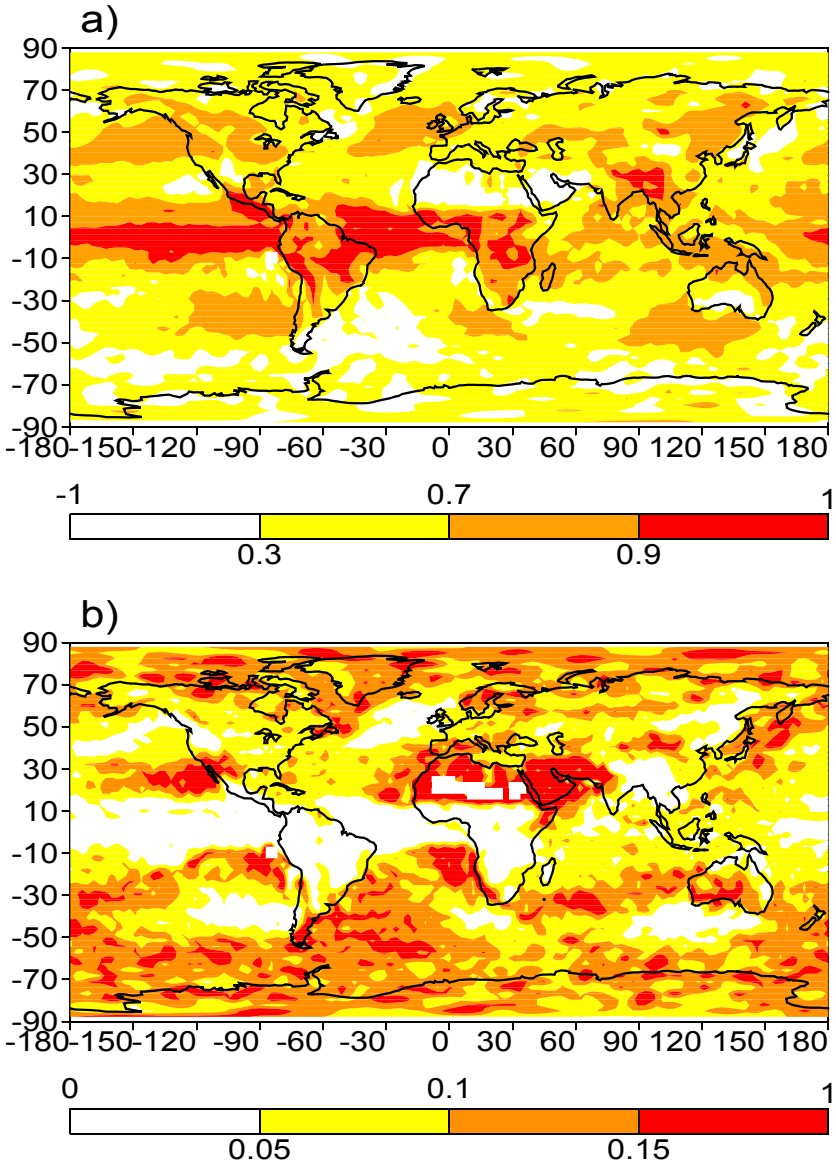


Figure 5: Maps of a) the ensemble mean of zero-lag temporal cross-correlations $\langle r(i,j) \rangle$ for seasonal precipitation, and b) the corresponding intraensemble scatter $\Delta r(i,j)$, as simulated for the AMIP decade by the ECMWF cycle 36 model. The ensemble-mean statistics are determined from N=15 realizations of $r(i,j)$ for precipitation with the seasonal cycle included.

Figure 6a

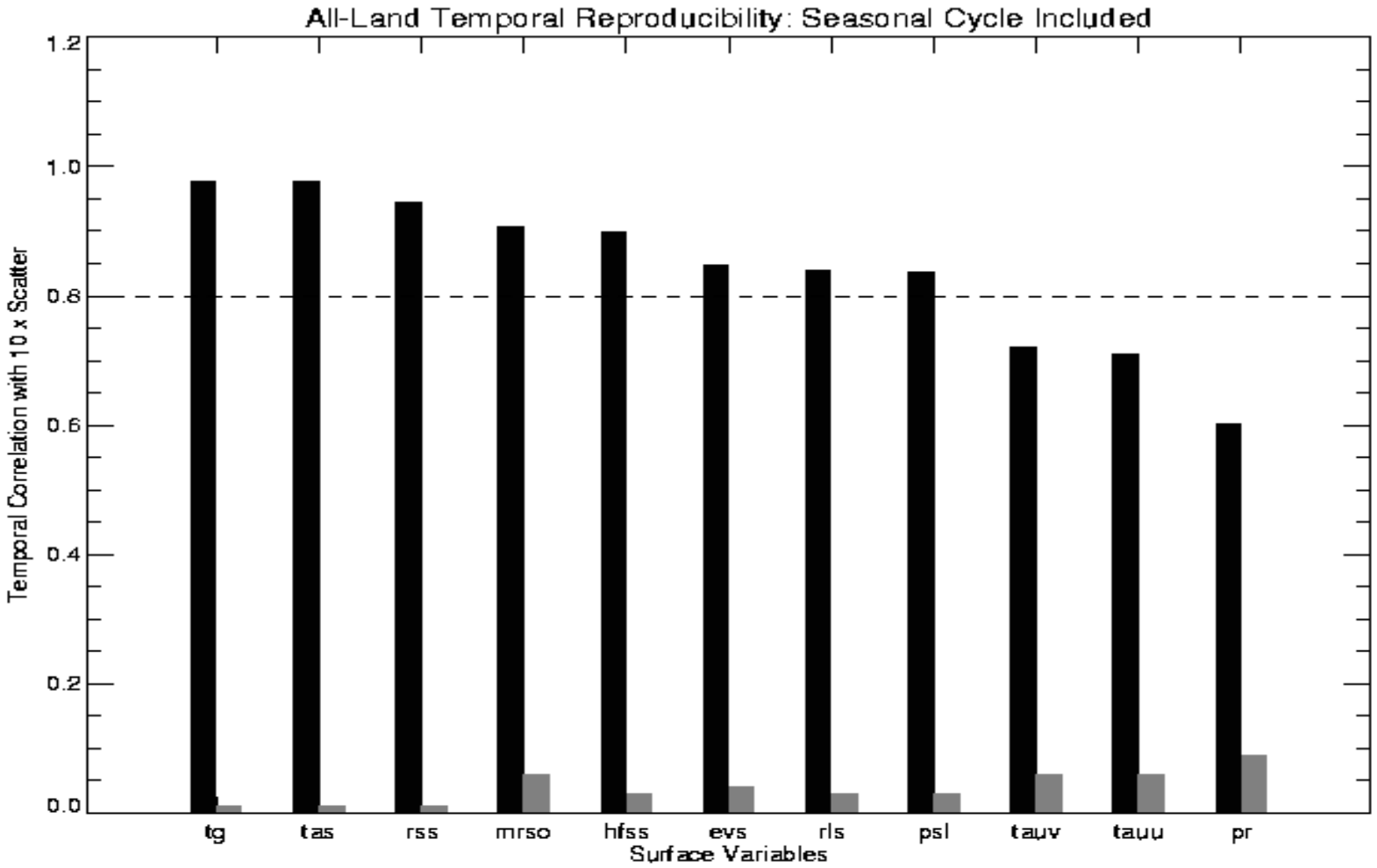


Figure 6: Global area-weighted land-averages <*R*> (dark bars) and intraensemble scatter Δ*R multiplied by a factor of 10 for visibility* (gray bars) of ensemble-mean zero-lag temporal correlations <*r*(i,j)> are shown for eleven ECMWF cycle 36 model surface variables (see Table 1 for acronym definitions). The ordering by model variable is according to descending values of <*R*>, and the dashed horizontal lines indicate an arbitrary common reference value of 0.8. In a) <*R*> and Δ*R* pertain to ensemble-mean correlations <*r*(i,j)> that are computed from seasonally averaged variables (i.e. from time series of MAM, JJA, SON, and DJF means).

Figure 6b

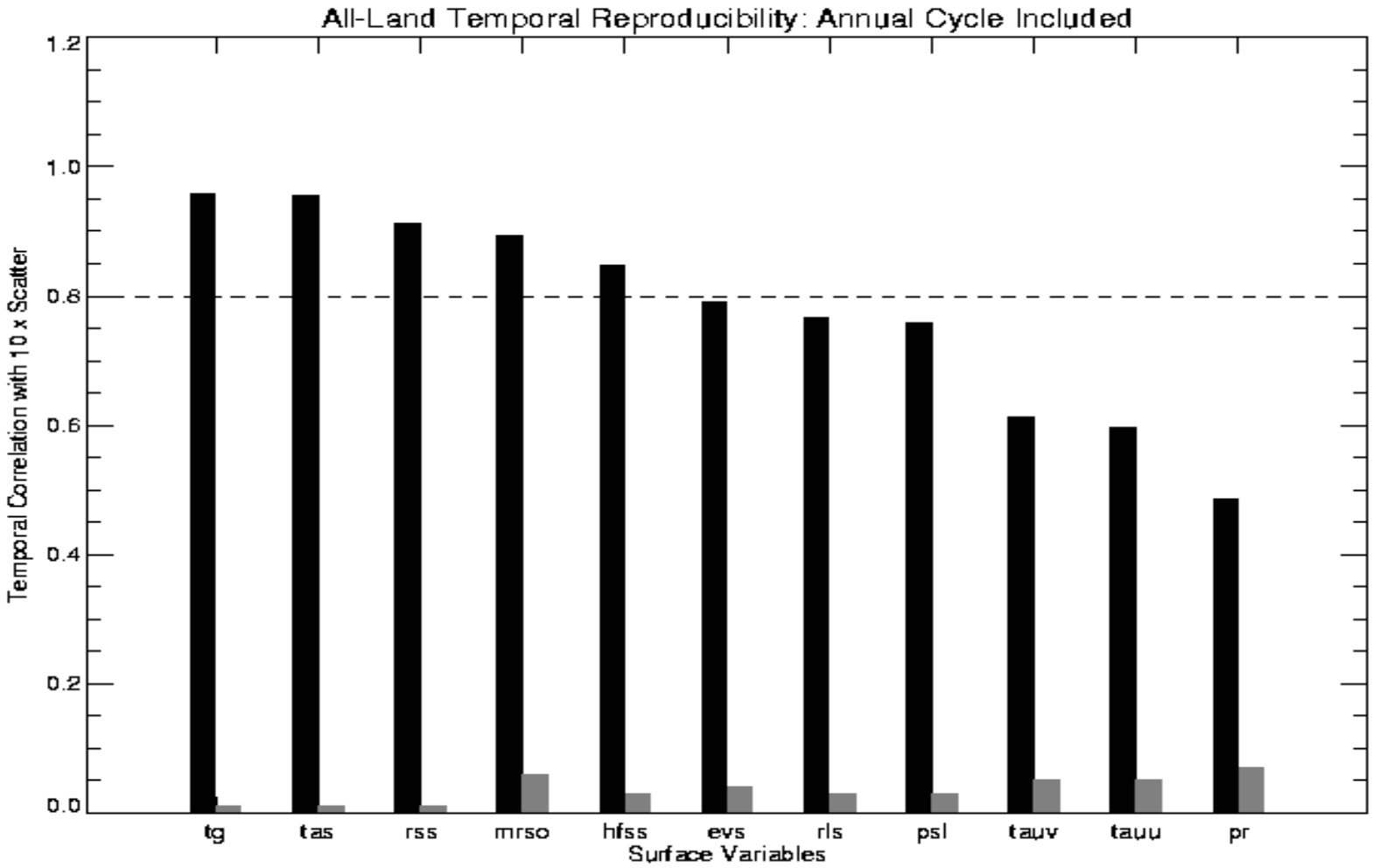


Figure 6b: As in 6a, except that the ensemble-mean correlations $\langle r(i,j) \rangle$ are calculated over monthly averaged variables.

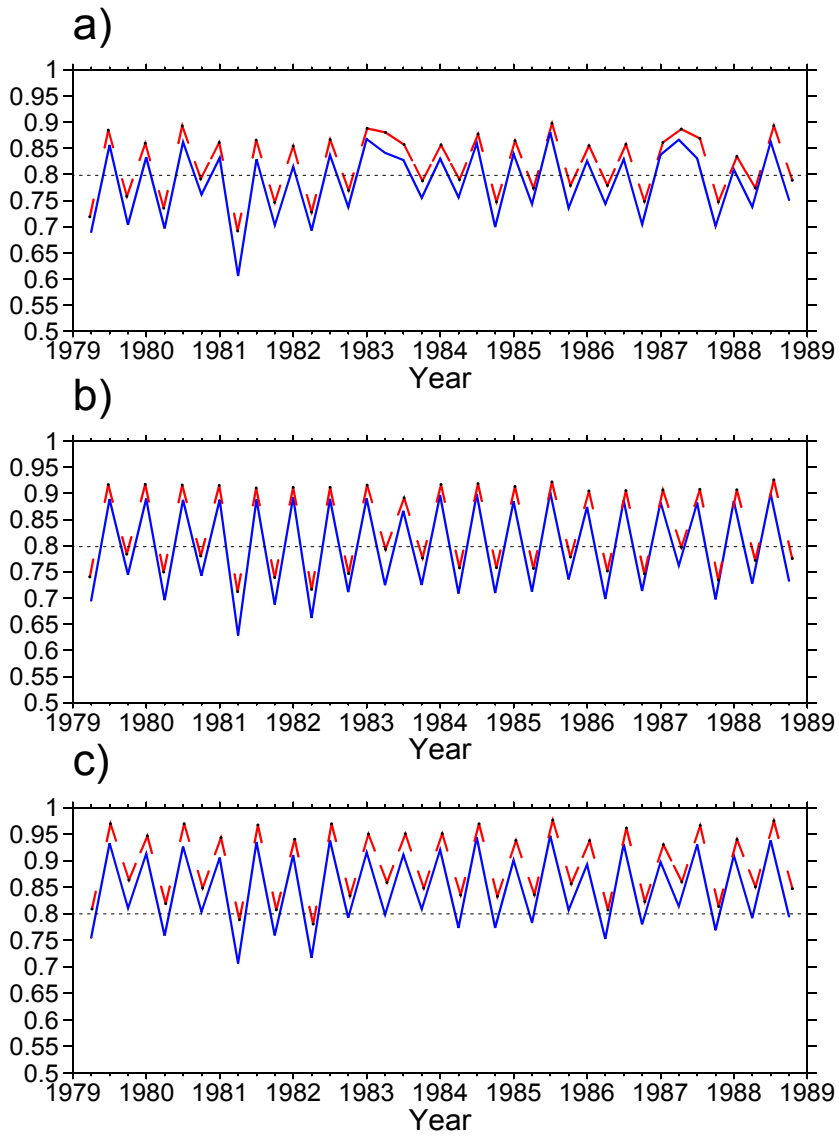


Figure 7: Time series of the ensemble mean of spatial pattern cross-correlations <*s*(k)> for seasonal precipitation as simulated for the decade 1979-1988 by the ECMWF cycle 36 model (solid lines), and with the positive-valued scatter |Δ*s*(k)| about <*s*(k)> added (broken lines**)**. As arbitrary common references, the dashed horizontal lines indicate spatial pattern correlation values of 0.8. In a) the pattern correlations are computed globally, in b) they are computed over continental points, and in c) only over tropical (30 S to 30 N) land points.

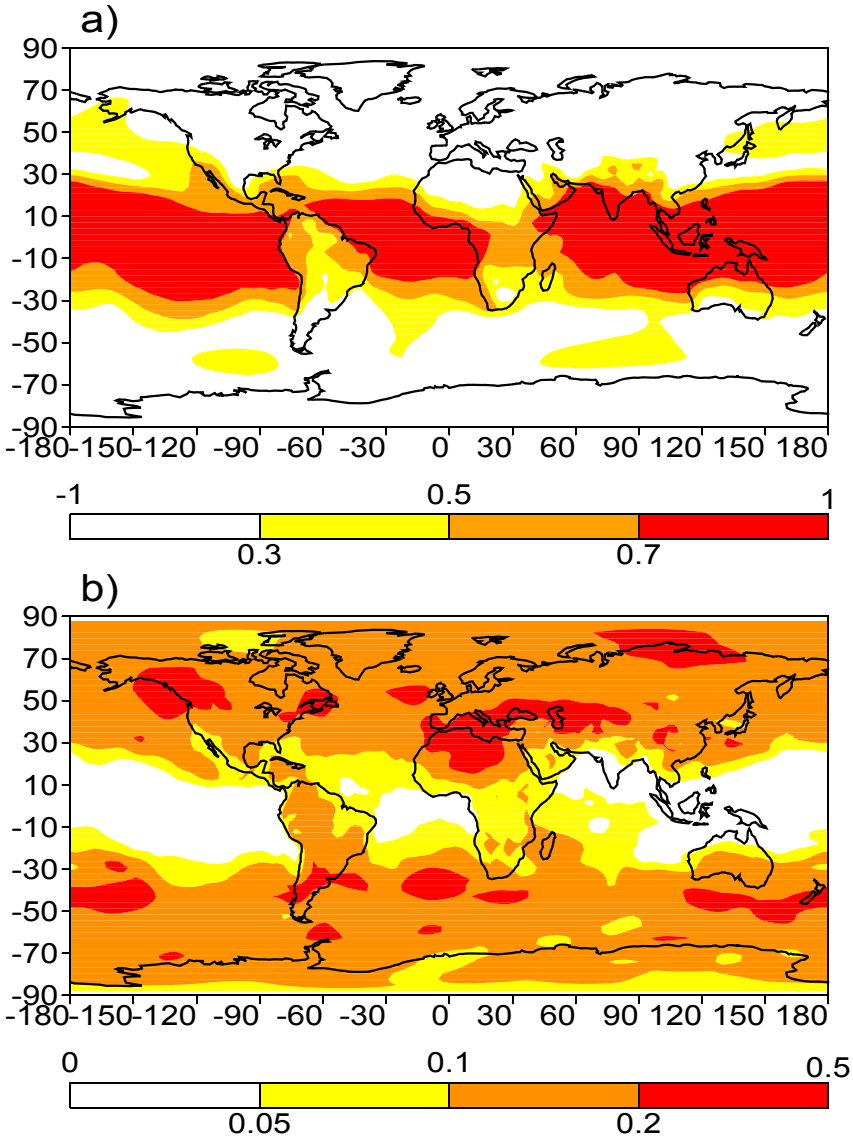


Figure 8: Maps of a) the ensemble mean of zero-lag temporal cross-correlations $\langle r(i,j) \rangle$ for seasonal anomalies (i.e. with the seasonal cycle removed) of MSLP, as simulated for the decade 1979-1988 by the ECMWF cycle 36 model, and of b) the intraensemble scatter $\Delta r(i,j)$ about $\langle r(i,j) \rangle$. The ensemble-mean statistics are determined from N = 15 realizations of $r(i,j)$.

Figure 9a

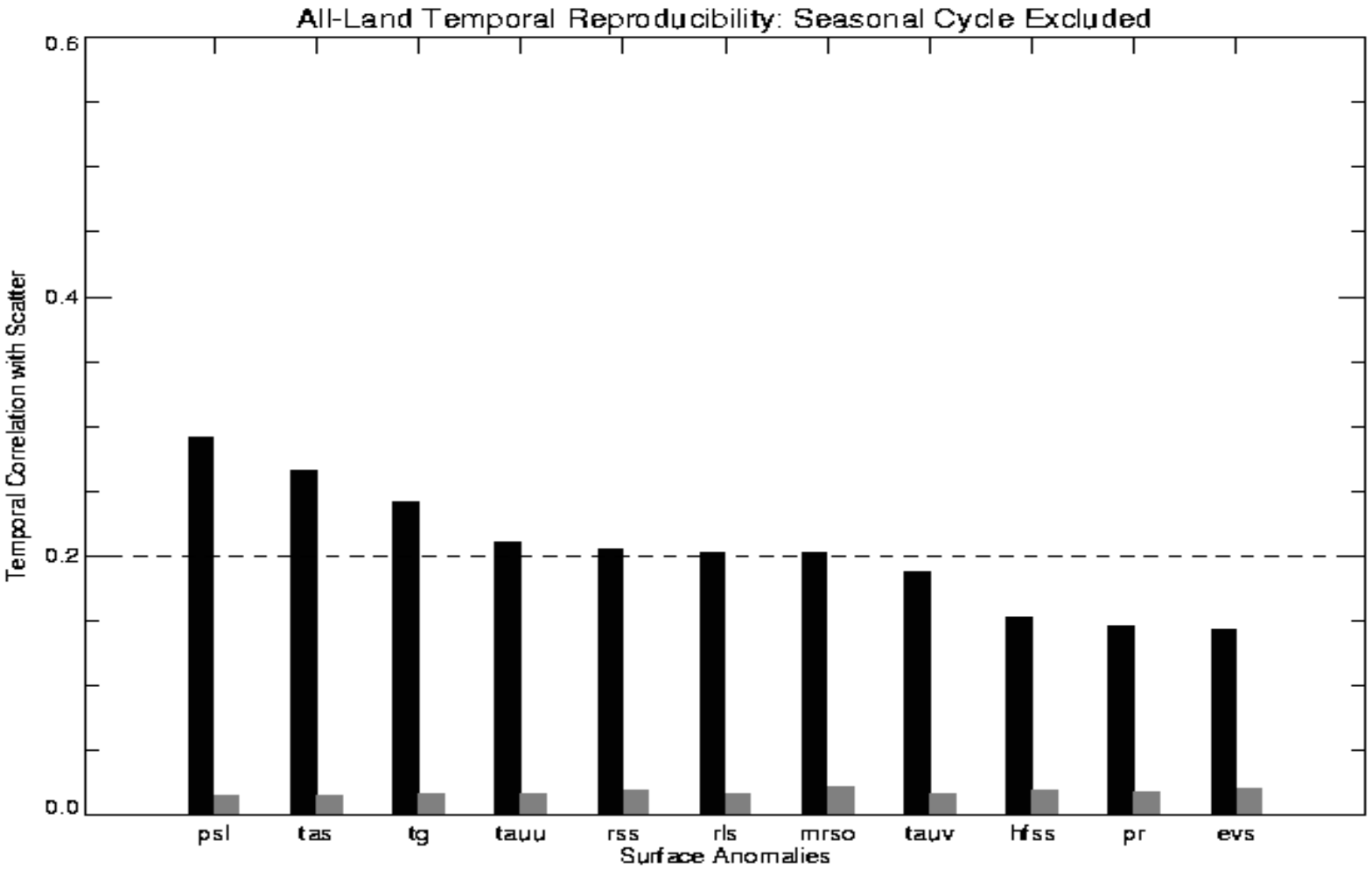


Figure 9: The ensemble mean <*R*> (dark bars) and intraensemble scatter Δ*R*(gray bars) of area-weighted land averages R of zero-lag temporal cross-correlations *r*(i,j) for seasonal anomalies of eleven ECMWF model surface variables (see Table 1 for key). The ensemble-mean statistics are computed from N=15 realizations, and are ordered by surface variable according to descending values of all-land <*R*>. The dashed horizontal lines indicate an arbitrary common reference value of 0.2. In a), the area averages are computed over all continental points.

Figure 9b

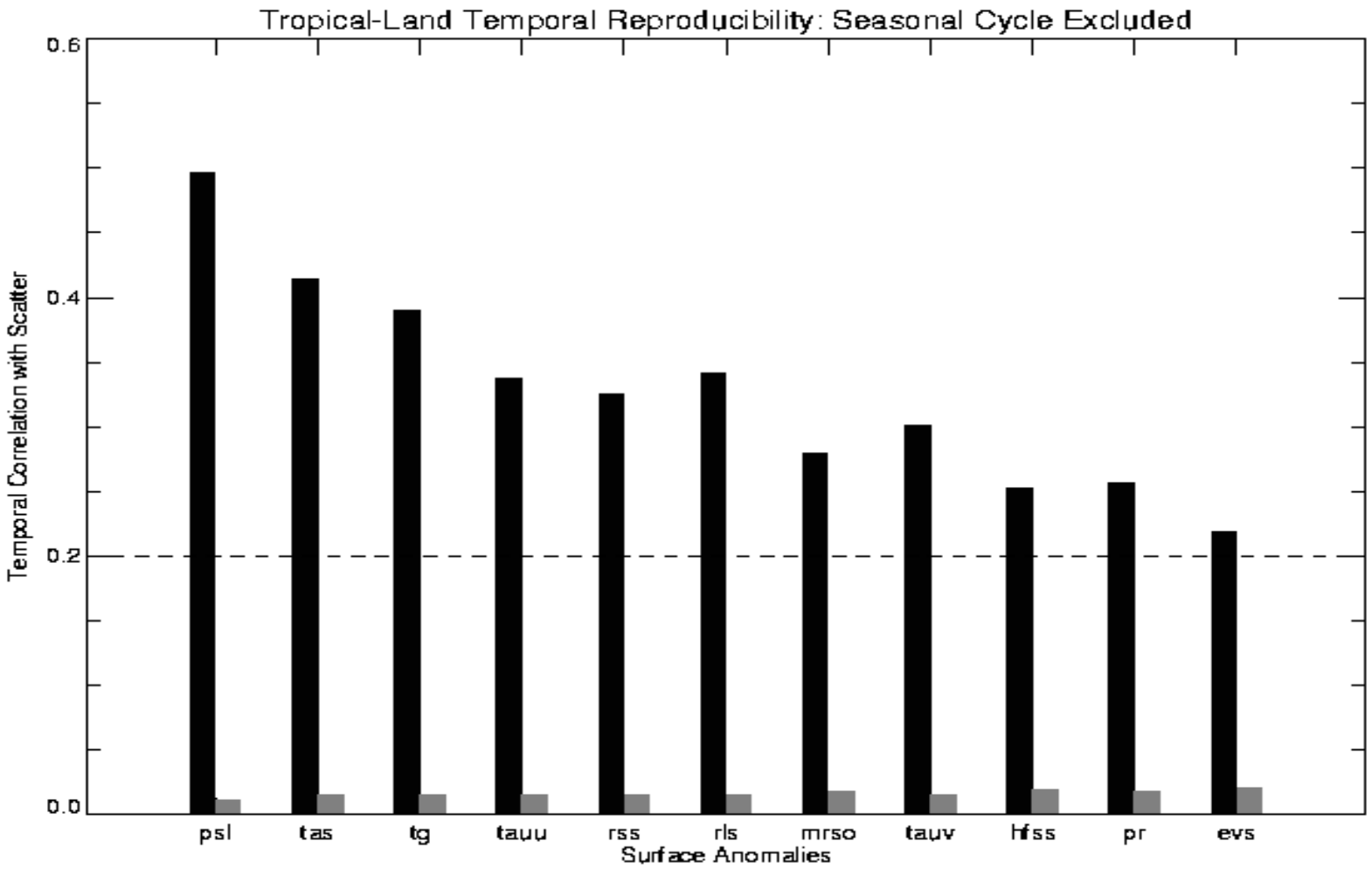


Figure 9b: As in 9a, except that the area averages are computed over only tropical (30S to 30N) land.

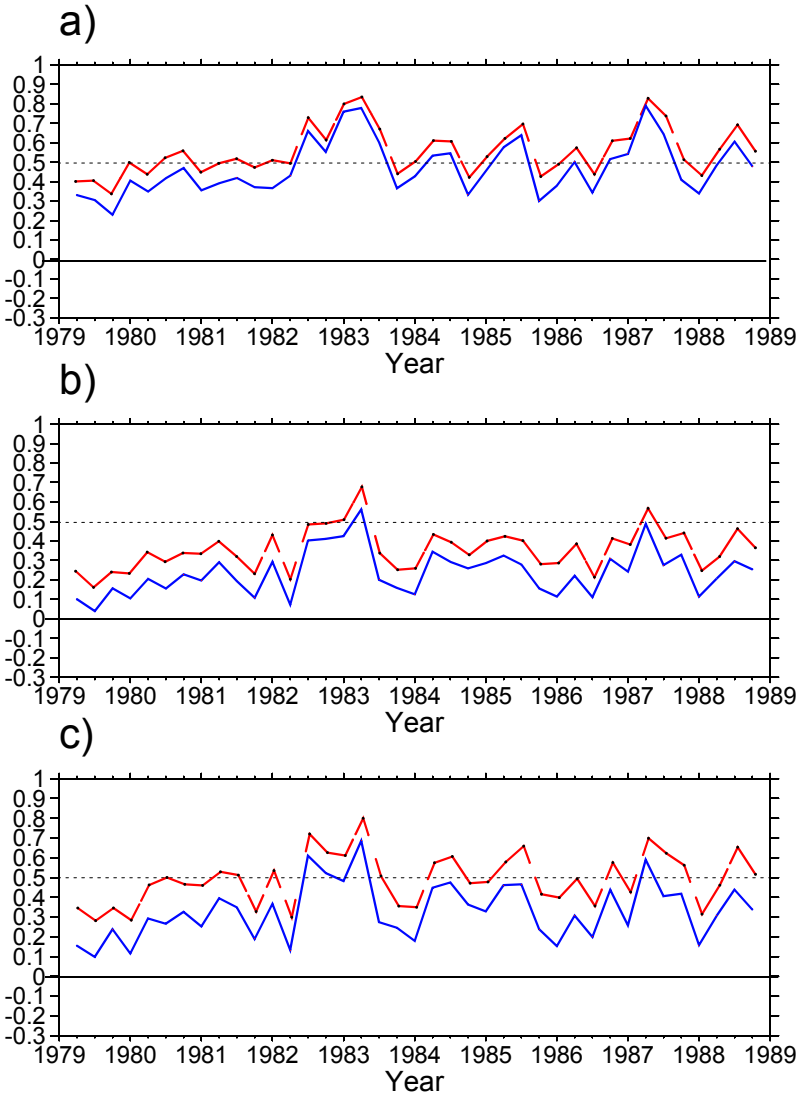


Figure 10: Time series of the ensemble mean of spatial pattern cross-correlations <*s*(k)> of seasonal anomalies of precipitation simulated by the ECMWF cycle 36 model for the decade 1979-1988 (solid line), and with the positive-valued intraensemble scatter |Δ*s*(k)| about <*s*(k)> added (broken line). The time series comprises 39 seasonal samples k, the first at MAM of 1979 and the last at SON of 1988. (The MAM points are situated just to the right of the major tick marks labeled by each year on the abscissa.) The horizontal lines indicate an arbitrary reference value of 0.5. In a), the pattern correlations are computed globally (both land and ocean points), in b) only over land points, and in c) only over tropical (30 S to 30 N) land points. In all cases, the ensemble-mean statistics are computed from N=15 realizations of the spatial pattern correlations *s*(k).

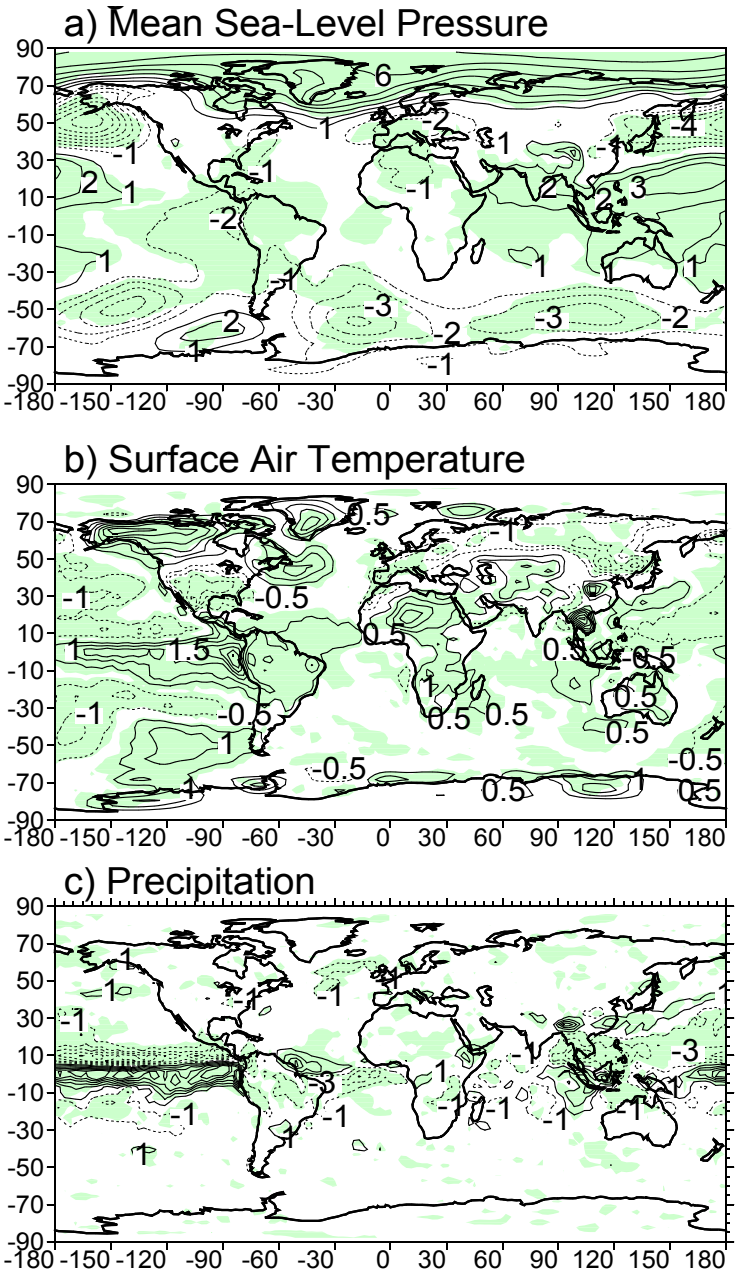


Figure 11: Maps of the average seasonal anomalies (isolines) of a) mean sea-level pressure (in hPa), of b) surface air temperature (in K), and of c) precipitation (in mm day^{-1}) computed across 6 realizations of the season MAM 1983 that coincided with an historically intense El Niño. The shading indicates where these mean anomalies are significantly different (at the 95% confidence level) from those in the MAM season of the other years of the decade 1979 to 1988 in which an El Niño did not occur (i.e., MAM of all years except 1987).

Figure 12a

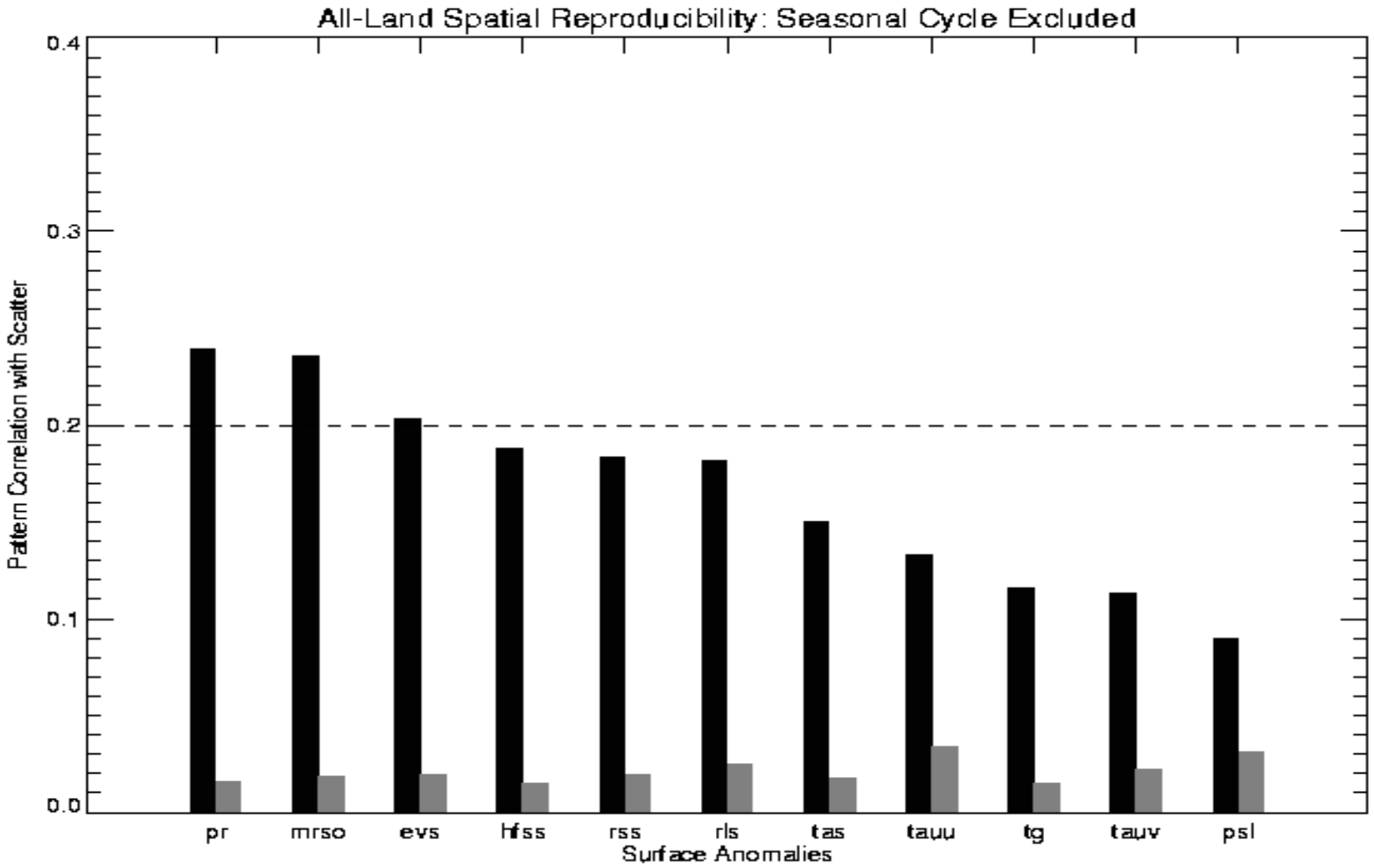


Figure 12: The ensemble mean <*S*>(dark bars) and scatter Δ*S* (gray bars) of the temporal average *S* of time series of pattern cross-correlations *s*(k), k = 1, ..., 39 for eleven ECMWF cycle 36 model seasonal surface anomalies (see Table 1 for key) are shown. The ensemble statistics are computed from N = 15 realizations of *s*(k), and are ordered by surface variable according to descending values of all-land <*S*>. The dashed horizontal lines indicate an arbitrary reference value of 0.2. In a) the ensemble-mean statistics are calculated for *s*(k) over all continental points.

Figure 12b

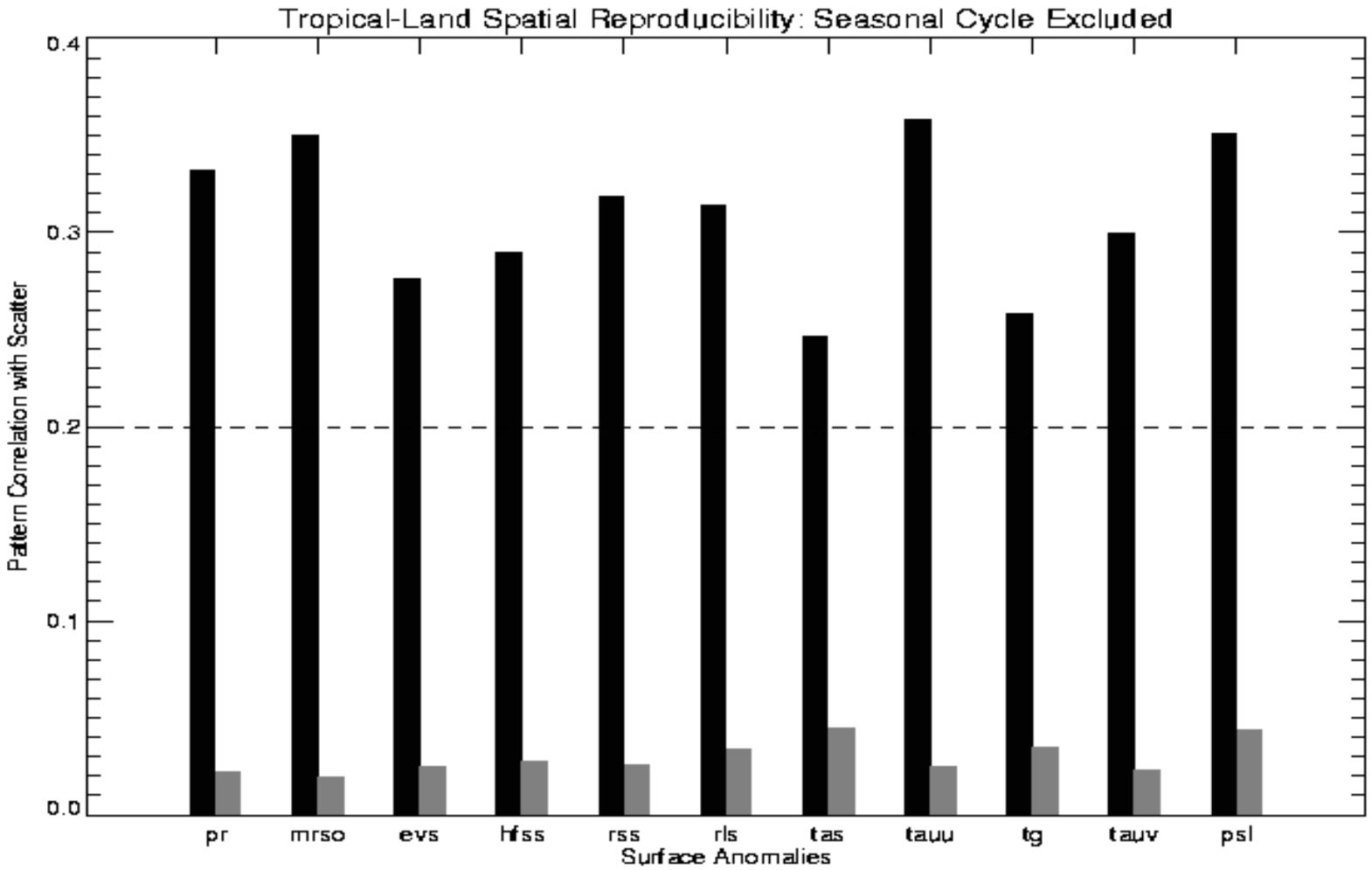


Figure 12b: As in 12a, except that the ensemble-mean statistics are calculated for s(k) only over tropical (30S to 30N) land.

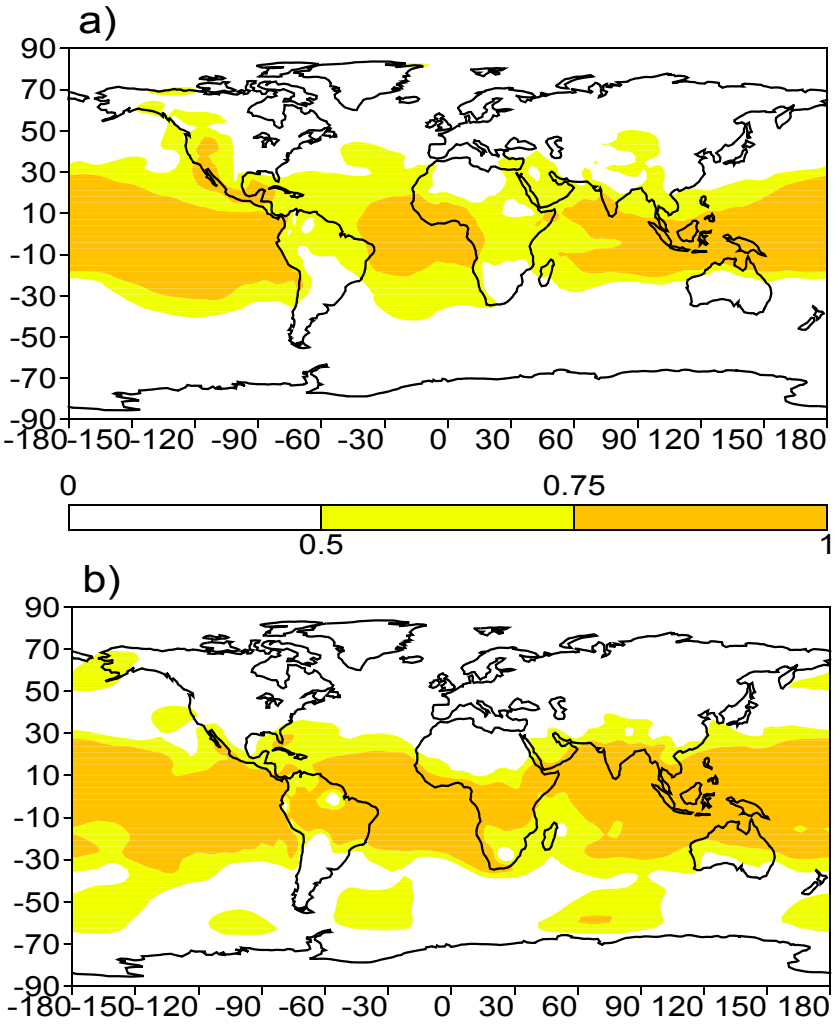


Figure 13: Geographical distribution of the fraction *f* of variance of mean sea-level pressure (MSLP) explained by the common ocean boundary forcing in seasonal means a) JJA and b) DJF of the decade 1979-1988, as simulated by the ECMWF cycle 36 model. Shaded areas $f > 0.5$ are indicative of a potentially predictable signal. Compare also with the temporal reproducibility metric for MSLP in Figure 7a.

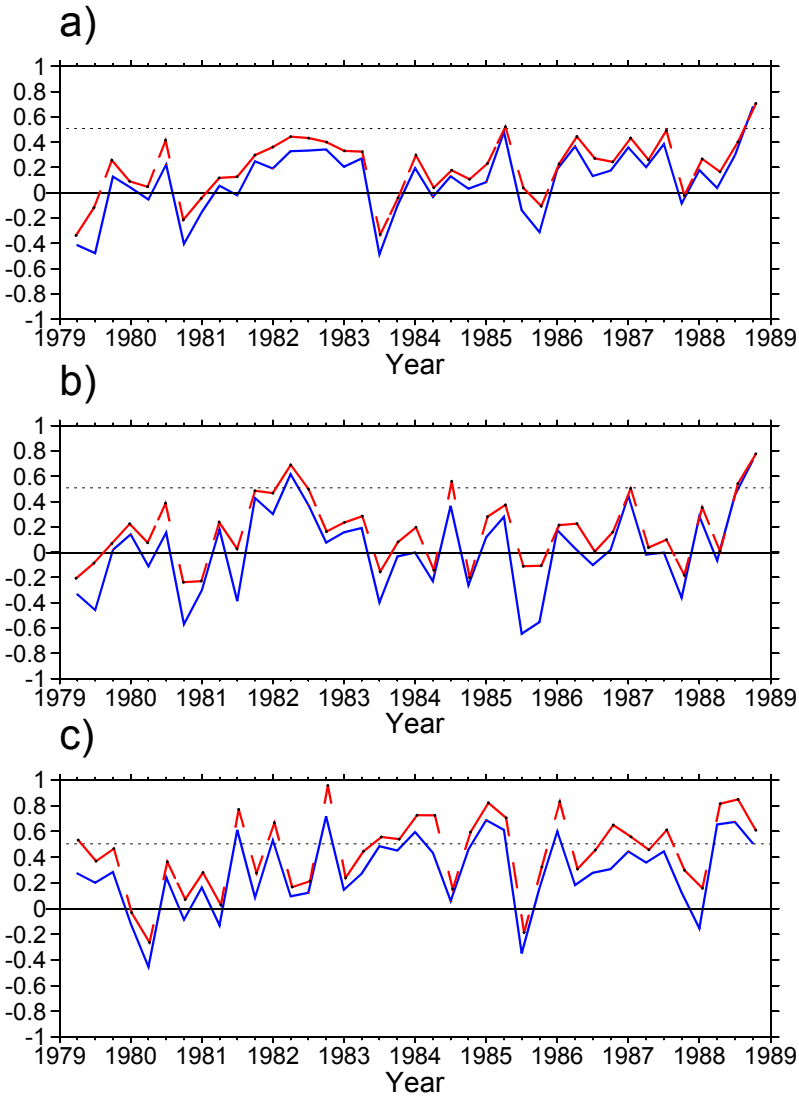


Figure 14: Time series of spatial pattern correlations *c*(k) of the ensemble-mean seasonal anomaly of simulated mean sea-level pressure against that from the NCEP-NCAR reanalysis data (solid line), and with the estimated uncertainty in *c*(k) also added (broken line). The horizontal dashed lines indicate a correlation of 0.5, which signifies a minimum level of practically useful forecast skill. In a) the spatial pattern correlations are computed globally, in b) over all land points, and in c) only over tropical (30 S to 30 N) land points.

Figure 15a

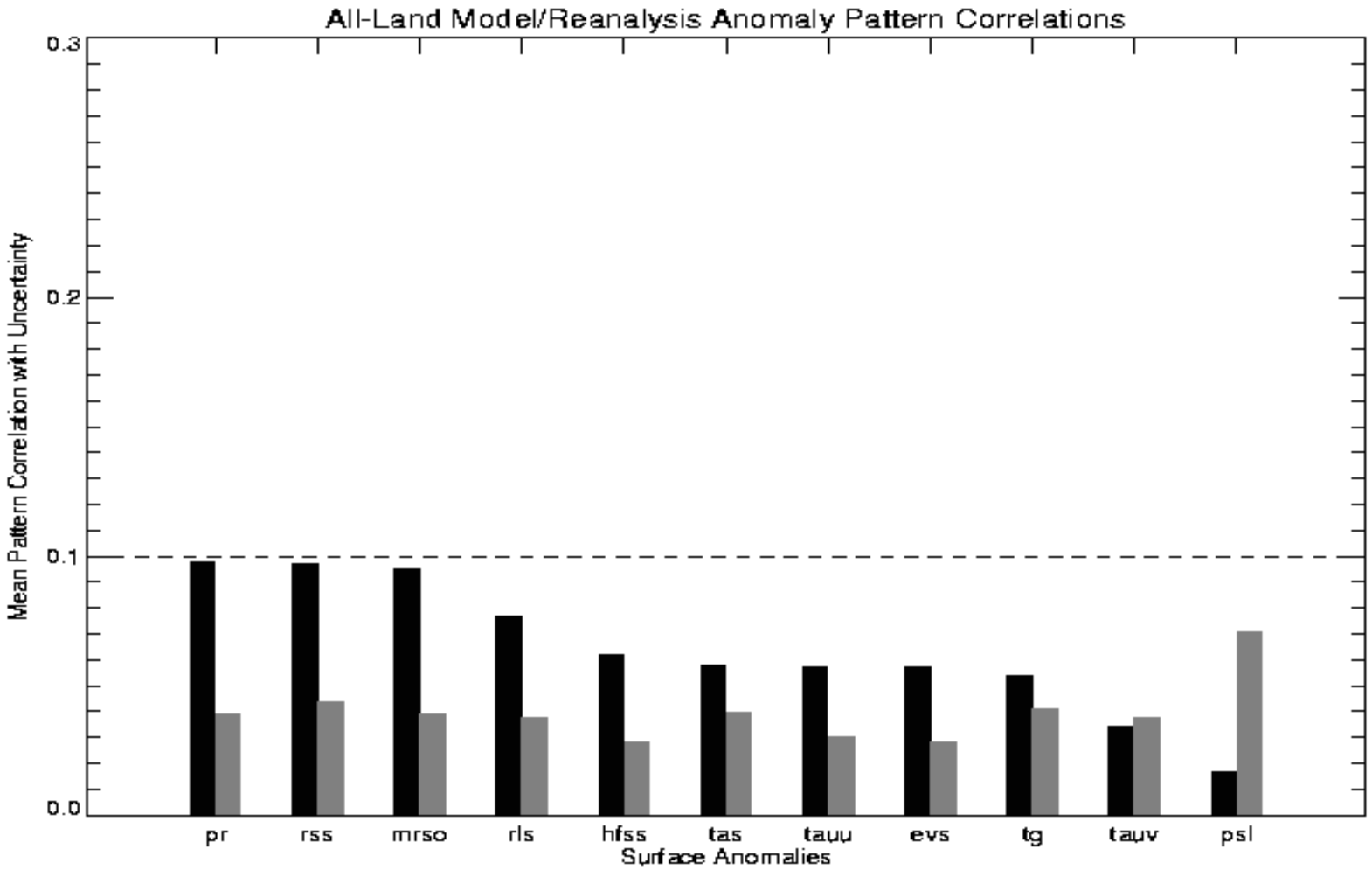


Figure 15: The temporal means *C* (dark bars) of the time series *c*(k), season k = 1,...,39 of pattern correlations of the ECMWF cycle 36 model ensemble-mean seasonal land-surface anomalies (see Table 1 for key) with corresponding anomalies from the NCEP-NCAR reanalysis, along with the estimated uncertainties (gray bars) are shown. The statistics are calculated from n = 6 model realizations and from a single realization of the NCEP-NCAR reanalysis, and these are ordered by surface variable according to descending all-land *C* values. The dashed horizontal lines indicate an arbitrary reference value of 0.1. In a) the *C* values are for correlations *c*(k) that are computed over all continental points.

Figure 15b

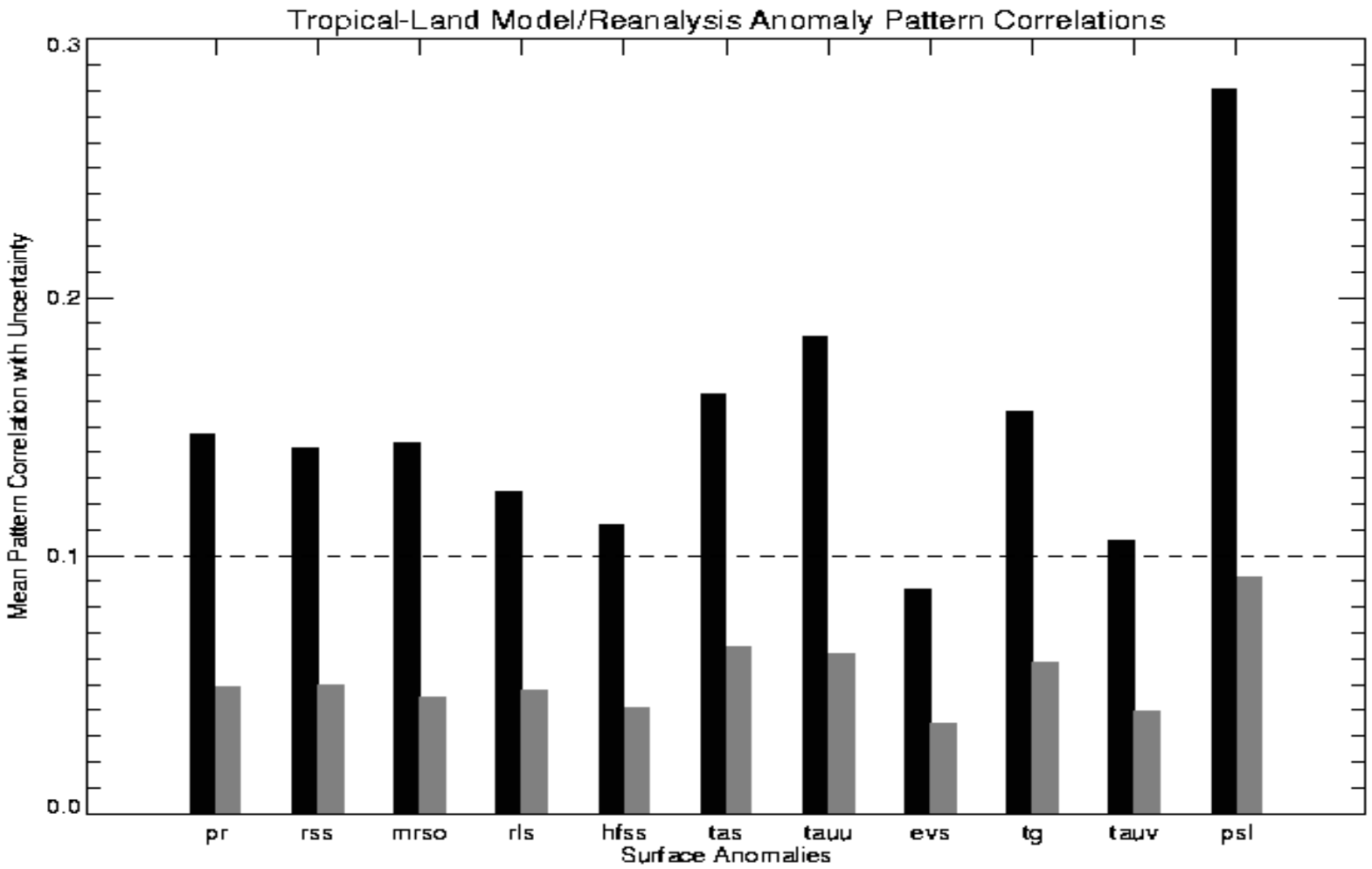


Figure 15b: As in 14 a, except that all statistics are computed only over tropical (30 S to 30 N) land.